

STIC-Biotech/ChemLib

119478

From: Hutzell, Paula
Sent: Thursday, April 15, 2004 11:11 AM
To: Graser, Jennifer; STIC-Biotech/ChemLib
Subject: RE: rush search

RECEIVED
APR 15 2004
STIC-BIOTECH/CHM LIB
(STIC)

approved

-----Original Message-----

From: Graser, Jennifer
Sent: Thursday, April 15, 2004 10:03 AM
To: Hutzell, Paula
Subject: rush search
Importance: High

Hi Paula,

Could you please authorize the following rush search for an amendment which is due?

Thanks,
Jennifer

STIC:

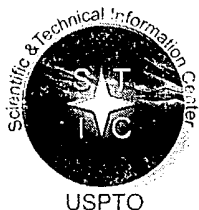
Please search SEQ ID NOS: 1; 1 (101-1176); 2; 3; 4; 5; 6; 6 (101-853); 6 (853-1176) from **10/608,504** in pending and commercial databases. They are species of one another and I need all of them searched.

Thanks,
Jennifer Graser
REMSEN
EO3 B09/ 3C18 mailbox
AU 1645
272-0858

Searcher: D. Schweiber
Phone: 272-2526
Location: Remsen EO1 A61
Date Picked Up: _____
Date Completed: 4/20
Searcher Prep/Review: 15
Clerical: _____
Online time: 10

TYPE OF SEARCH:
NA Sequences: 7
AA Sequences: 2
Structures: _____
Bibliographic: _____
Litigation: _____
Full text: _____
Patent Family: _____
Other: _____

VENDOR/COST (where applic.)
STN: _____
DIALOG: _____
Questel/Orbit: _____
DRLink: _____
Lexis/Nexis: _____
Sequence Sys.: Compage
WWW/Internet: _____
Other (specify): _____



STIC SEARCH RESULTS FEEDBACK FORM

Biotech-Chem Library

Questions about the scope or the results of the search? Contact *the searcher or contact:*

Mary Hale, Information Branch Supervisor
Remsen Bldg. 01 D86
571-272-2507

Voluntary Results Feedback Form

➤ I am an examiner in Workgroup: Example: 1610

➤ Relevant prior art **found**, search results used as follows:

- ☐ 102 rejection
- ☐ 103 rejection
- ☐ Cited as being of interest.
- ☐ Helped examiner better understand the invention.
- ☐ Helped examiner better understand the state of the art in their technology.

Types of relevant prior art found:

- ☐ Foreign Patent(s)
- ☐ Non-Patent Literature
(journal articles, conference proceedings, new product announcements etc.)

➤ Relevant prior art **not found**:

- ☐ Results verified the lack of relevant prior art (helped determine patentability).
- ☐ Results were not useful in determining patentability or understanding the invention.

Comments:

Drop off or send completed forms to STIC-Biotech-Chem Library Remsen Bldg.



GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 17, 2004, 06:48:18 ; Search time 415.78 Seconds
(without alignments)
9221.792 Million cell updates/sec

Title: US-10-608-504-1

Perfect score: 1271

Sequence: 1 gcgcgcataacgaatcctag.....agcctaacgacagcgcct 1271

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database :
1: em_estba:*
2: em_esthum:*
3: em_estmu:*
4: em_estov:*
5: em_estov:*
6: em_estov:*
7: em_estro:*
8: em_estro:*
9: gb_est1:*
10: gb_est2:*
11: gb_est3:*
12: gb_est4:*
13: gb_est5:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_esthum:*
18: em_estinv:*
19: em_estpin:*
20: em_estvrt:*
21: em_estvrt:*
22: em_estvrt:*
23: em_estvrt:*
24: em_estvrt:*
25: em_estvrt:*
26: em_estvrt:*
27: em_estvrt:*
28: gb_est1:*
29: gb_est2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	54.6	4.3	473	28	AO936633	AO936633 HSJ1-905
2	45	3.5	1201	13	BX356664	BX356664 BX356664
3	41.8	3.3	1201	13	BX376097	BX376097 BX376097
4	40.8	3.2	1201	13	BX381961	BX381961 BX381961

RESULT 1	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	MEDLINE	PUBMED	COMMENT
AO936633	HSJ1-905	Human NotI clones Homo sapiens genomic, genomic survey	AO936633	1	GI:7213011	GSS.	Homo sapiens (human)	1 (bases 1 to 473)	Zabarovskiy, E.R., Gizaullin, R., Podowski, R.M., Zabarovskaya, V.V., Xie, L., Muravenko, O.V., Kozhnev, S., Petrenko, L., Skobeleva, N., Li, J., Protodopov, A., Kashub, V., Ernberg, I., Winberg, G. and Wahlstedt, C.	NotI clones in the analysis of the human genome	Nucleic Acids Res.	28 (7), 1635-1639 (2000)	Contact: Podowski RM Center for Genomics Research Karolinska Institute 17177 Stockholm, Sweden Tel: +46-8-728-6372 Fax: +46-8-337953	

ALIGNMENTS

5	39.4	3.1	692	13	BX540413
C	6	39.4	3.1	718	9
C	7	39.4	3.1	743	29
C	8	39.4	3.1	756	13
C	9	39.4	3.1	787	29
C	10	39.4	3.1	885	13
C	11	39.4	3.1	995	29
C	12	39.4	3.1	959	10
C	13	39.4	3.1	959	29
C	14	38.6	3.0	1201	13
C	15	38.6	3.0	835	14
C	16	37.6	3.0	997	29
C	17	37.4	2.9	521	9
C	18	37.4	2.9	697	28
C	19	37.4	2.9	1108	13
C	20	37.4	2.9	1201	9
C	21	37.2	2.9	423	28
C	22	37.2	2.9	730	12
C	23	37.2	2.9	758	29
C	24	37.2	2.9	1454	10
C	25	37.2	2.9	535	9
C	26	37.2	2.9	603	9
C	27	37.2	2.9	604	9
C	28	37.2	2.9	608	9
C	29	37.2	2.9	613	9
C	30	37.2	2.9	676	12
C	31	37.2	2.9	681	12
C	32	36.8	2.9	464	28
C	33	36.8	2.9	469	12
C	34	36.8	2.9	1201	29
C	35	36.6	2.9	539	10
C	36	36.6	2.9	942	29
C	37	36.6	2.9	1134	13
C	38	36.4	2.9	298	12
C	39	36.4	2.9	910	28
C	40	36.2	2.8	673	12
C	41	36.2	2.8	827	29
C	42	36.2	2.8	842	13
C	43	36.2	2.8	902	29
C	44	36.2	2.8	962	13
C	45	36.2	2.8	1101	13

Db	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
Db	BX381961	1201 bp linear EST 08-MAY-2003	BX381961									
Qy	BX381961	1201 bp linear EST 08-MAY-2003	BX381961									
Db	BX381961	1201 bp linear EST 08-MAY-2003	BX381961									
Qy	BX381961	1201 bp linear EST 08-MAY-2003	BX381961									
Db	BX381961	1201 bp linear EST 08-MAY-2003	BX381961									
Qy	BX381961	1201 bp linear EST 08-MAY-2003	BX381961									
Db	BX381961	1201 bp linear EST 08-MAY-2003	BX381961									
Qy	BX381961	1201 bp linear EST 08-MAY-2003	BX381961									
Db	BX381961	1201 bp linear EST 08-MAY-2003	BX381961									
Qy	BX381961	1201 bp linear EST 08-MAY-2003	BX381961									
Db	BX381961	1201 bp linear EST 08-MAY-2003	BX381961									
Qy	BX381961	1201 bp linear EST 08-MAY-2003	BX381961									
Db	BX381961	1201 bp linear EST 08-MAY-2003	BX381961									
Qy	BX381961	1201 bp linear EST 08-MAY-2003	BX381961									
Db	BX381961	1201 bp linear EST 08-MAY-2003	BX381961									
Qy	BX381961	1201 bp linear EST 08-MAY-2003	BX381961									
Db	BX381961	1201 bp linear EST 08-MAY-2003	BX381961									
Qy	BX381961	1201 bp linear EST 08-MAY-2003	BX381961									
Db	BX381961	1201 bp linear EST 08-MAY-2003	BX381961									
Qy	BX381961	1201 bp linear EST 08-MAY-2003	BX381961									
Db	BX381961	1201 bp linear EST 08-MAY-2003	BX381961									
Qy	BX381961	1201 bp linear EST 08-MAY-2003	BX381961									
Db	BX381961	1201 bp linear EST 08-MAY-2003	BX381961									
Qy	BX381961	1201 bp linear EST 08-MAY-2003	BX381961									
Db	BX381961	1201 bp linear EST 08-MAY-2003	BX381961									
Qy	BX381961	1201 bp linear EST 08-MAY-2003	BX381961									
Db	BX381961	1201 bp linear EST 08-MAY-2003	BX381961									
Qy	BX381961	1201 bp linear EST 08-MAY-2003	BX381961									
Db	BX381961	1201 bp linear EST 08-MAY-2003	BX381961									
Qy	BX381961	1201 bp linear EST 08-MAY-2003	BX381961									
Db	BX381961	1201 bp linear EST 08-MAY-2003	BX381961									
Qy	BX381961	1201 bp linear EST 08-MAY-2003	BX381961									
Db	BX381961	1201 bp linear EST 08-MAY-2003	BX381961									
Qy	BX381961	1201 bp linear EST 08-MAY-2003	BX381961									
Db	BX381961	1201 bp linear EST 08-MAY-2003	BX381961									
Qy	BX381961	1201 bp linear EST 08-MAY-2003	BX381961									
Db	BX381961	1201 bp linear EST 08-MAY-2003	BX381961									

[illegible]

Fax: +33 (0) 2.23.48.54.70

Email: Madeleine.Douaïre@roazhon.inra.fr
 Sequence cleaned of vector, adaptor and repetitions. Contact us
 at sigenaport@jouy.inra.fr to obtain the chromatogram of this
 sequence.

Plate: 0001 row: d column: 10
 Seq primer: M13R

FEATURES

source

Location/Qualifiers

1..692

/organism="Gallus gallus"

/mol_type="mRNA"

/db_xref="taxon:9031"

/clone="gscab0001b.d.10"

/tissue_type="adipose tissue, granulosa, multi-tissues,

oviduct, small follicle, utero-vaginal gland"

/dev_stage="adult"

/lab_host="DH10B"

/clone_1ib="AGENAE Gallus gallus multi-tissues library

/note="Vector: pTTT3D-pac; Clone distribution: AGENAE

Resource centre: Francois Piumi, Francois Piumi.inra.fr,

INRA, CEA Radiobiologie et Etude du genome (LRBG), Domaine

de Vilvert, 78352, Jouy-en-Josas cedex, FRANCE"

ORIGIN

Query Match 3.1%; Score 39.4; DB 13; Length 692;

Best Local Similarity 62.9%; Pred. No. 10; Mismatches 36; Indels 0; Gaps 0;

Matches 61; Conservative 0; Mismatches 36; Indels 0; Gaps 0;

QY 805 CTGGGCTCTGTGACCATCCGCTACTTCTTCTGGGAAGCTGCTAATGCACTGAT 864

DB 321 CTGGGCGAGTTGGCCATCTGCTCTGCACTTTSCACACAGCAGAGTACCGCAT 262

QY 865 TTCTCCTGATTCCTCTGTTGTCGCAATGTCAG 901

DB 261 TTTCGTAAGTCACCTCTTCTGTCAGCCTCTGAG 225

RESULT 6

AJ397466/c 718 bp mRNA linear EST 25-JAN-2001

DEFINITION AJ397466 dkfz426 Gallus gallus cDNA clone 217r1, mRNA sequence.

ACCESSION AJ397466

VERSION AJ397466.1 GI:7130433

KEYWORDS EST.

SOURCE Gallus gallus (chicken)

ORGANISM Gallus gallus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Archosauria; Aves; Neognathae; Galliformes; Phasianidae;

Phasianinae; Gallus.

1 (bases 1 to 718)

Plachy J., Korn B., and Buerstedde J.M.

A large database of chickenursal ESTs as a resource for the

analysis of vertebrate gene function

Genome Res. 10 (12), 2062-2069 (2000)

JOURNAL MEDLINE PUBMED

COMMENT Contact: Buerstedde JM

Cellular Immunology

Heinrich-Pette-Institute

Martinstr. 52, 20251 Hamburg, Germany

Email: URL: http://genefix.hpi.uni-hamburg.de/dt0est.html.

Location/Qualifiers

1..718

/organism="Gallus gallus"

/mol_type="mRNA"

/strain="CB"

/db_xref="taxon:9031"

/clone="217r1"

/tissue_type="Bursa of Fabricius"

/cell_type="Bursal lymphocyte"

/dev_stage="2-3 weeks old"

ORIGIN

/clone_1ib="dkfz426"

Query Match 3.1%; Score 39.4; DB 9; Length 718;

Best Local Similarity 62.9%; Pred. No. 10;

Matches 61; Conservative 0; Mismatches 36; Indels 0; Gaps 0;

QY 805 CTGGGCTCTGTGACCATCCGCTACTTCTTCTGGGAAGCTGCTAATGCACTGAT 864

DB 117 CTGGGCGAGTTGGCCATCTGCTCTGCACTTTSCACACAGCAGAGTACCGCAT 58

QY 865 TTCTCCTGATTCCTCTGTTGTCGCAATGTCAG 901

DB 57 TTTCGTAAGTCACCTCTTCTGTCAGCCTCTGAG 21

RESULT 7

CC684922 743 bp DNA linear GSS 19-JUN-2003

LOCUS CC684922

DEFINITION CC684922 ZM 0.7-1.5 KB Zea mays genomic clone ZMMBMA0397J14,

genomic survey sequence.

ACCESSION CC684922

VERSION CC684922.1 GI:32089698

KEYWORDS GSS.

SOURCE Zea mays

ORGANISM Zea mays

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACAD

clade; Panicoideae; Andropogoneae; Zea.

1 (bases 1 to 743)

Whitefaw,C.A., Queckenbush,J., Van Aken,S., Uteback,T.,

Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,

Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.

Consortium for Maize Genomics

Unpublished (2002)

Other_GSS: OGRX55TV

Contact: Cathy Whitefaw

TIGR

9712 Medical Center Drive, Rockville, MD 20850, USA

Tel: 301-838-5843

Fax: 301-838-0208

Email: whitefaw@tigr.org

Seq primer: TR

Class: sheared ends.

Location/Qualifiers

1..743

/organism="Zea mays"

/mol_type="genomic DNA"

/strain="B73"

/db_xref="taxon:4577"

/clone="ZMMBMA0397J14"

/clone_1ib="ZM_0.7-1.5 KB"

/note="Vector: pBSK-; Site 1: HincII, 0.7-1.5 kb

methylation filtered genomic DNA library"

ORIGIN

Query Match 3.1%; Score 39.4; DB 29; Length 743;

Best Local Similarity 53.6%; Pred. No. 11;

Matches 82; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

QY 531 CCGCAGCTGTGTCGCGTGAATTATCTCAATGCAATAGCTTCACTCTACTGGG 590

DB 475 CTTCCGCTCCGCGCGTTCCTGTCCTCCACCTCTCATCGGATGGCAATTCATAG 534

QY 591 TATTCGCGGCTTCACCGGATGGCGCATGCGCAAGTGAATCTTTGAATTAAGGCC 650

DB 535 GTTTCGCGGCGGCGGATCCGATCCAGTCCGCTTCGCGAGCGCTTAAGCTTCAGGC 594

QY 651 TCGAGTTCGCGCTTGTGCTCTCTCTTTGTACGC 683

DB 595 ACGAGAGGCGGCTTGTGCTCTCTCTCTCTCTCCGTCGCCG 627

RESULT 8
BU448643/c
LOCUS
DEFINITION 60321136F1 CSEGRBN13 Gallus gallus cDNA clone CHEST191c7 5', mRNA

756 bp mRNA linear EST 29-NOV-2002

ACCESSION BU448643
VERSION BU448643
KEYWORDS EST
SOURCE Gallus gallus (chicken)
ORGANISM Gallus gallus

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.
1 (bases 1 to 756)
Boardman, P.E., Sant-Ezquerro, J., Overton, I.M., Burt, D.W., Bosch, E., Fong, W.T., Tickle, C., Brown, W.R.A., Wilson, S.A. and Hubbard, S.J.
A Comprehensive Collection of Chicken CDNA
Curr. Biol. 12 (22), 1965-1969 (2002)
2235534
1243532

TITLE JOURNAL
MEDLINE
PUBMED

COMMENT
Contact: Simon Hubbard
Department of Biomolecular Sciences
University of Manchester Institute of Science and Technology (UMIST)
PO Box 88, Manchester, M60 1QD, UK
Tel: 0161208930
Fax: 01612360409
Email: Simon.Hubbard@umist.ac.uk.
Location/Qualifiers
1. 756
/organism="Gallus gallus"
/mol_type="mRNA"
/strain="Layer"
/db_xref="taxon:9031"
/clone="CHEST191c7"
/sex="Female"
/dev_stage="adult"
/lab_host="MDH108"
/clone_lib="CSEGRBN13"
/note="Organ: ovary; Vector: Bluescript II KS(+); Site: 1: EcoRI; Site 2: NotI; This normalized library was constructed from 1 million independent clones. cDNA synthesis was initiated using an oligo(dt) primer, using methylated C in the first strand synthesis reaction. Following this first strand reaction, double-stranded cDNA was blunt-ended, ligated to NotI adapters, digested with EcoRI, size-selected, and cloned into the NotI and EcoRI compatible sites of a custom modified MCS of the Bluescript (KS+) vector. The library was normalized in 2 rounds using conditions adapted from Soares et al., PNAS (1996) 91: 9228-9232 and Bonaldo et al., Genome Research 6 (1996): 791, except that a significantly longer reannealing hybridization was used."

ORIGIN
Query Match 3.1%; Score 39.4; DB 13; Length 756;
Best Local Similarity 62.9%; Pred. No. 11;
Matches 61; Conservative 0; Mismatches 36; Indels 0; Gaps 0;

QY 805 CTTGGGCTCTGACACCCGCTACTCTCTTGGGAAGGCTGTAATGACACATAT 864
DB 270 CTTGGGCGAGTGGCGCATCTGCTCTGACCTTTCACACAGCAGACGTGACCCGAT 211
QY 865 TTCTCCTGATATCTCTTGTGTCGACATATGTGAG 901
DB 210 TTTTGTAAAGTCACCTTCTTGCTGAGCCTCTGAG 174

RESULT 9
CNS009WA/c
LOCUS
DEFINITION CNS009WA 787 bp DNA linear GSS 03-JUN-1999
Drosophila melanogaster genome survey sequence T7 end of BAC # BAC20K04 of RPCI-98 library from Drosophila melanogaster (fruit

fly), genomic survey sequence.
ACCESSION AL054013
VERSION AL054013.1 GI:4935590
KEYWORDS GSS.
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Preygora; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 787)

REFERENCE
AUTHORS Direct Submission
TITLE Genoscope.
JOURNAL Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage : BP 191 91006 Evry cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see <http://www.fruitfly.org> The BDGP Drosophila melanogaster BAC library was prepared by Kazuo Oosawa and Aaron Mammoser in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain Y2; cn bw sp, the same strain used for the BDGP's p1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
Location/Qualifiers
1. 787
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACR20K04"
/clone_lib="RPCI-98"
/note="end : 17"

ORIGIN
Query Match 3.1%; Score 39.4; DB 29; Length 787;
Best Local Similarity 45.5%; Pred. No. 11;
Matches 76; Conservative 15; Mismatches 76; Indels 0; Gaps 0;

QY 652 CGAGTGGCCCTTGTCTCTCTTGTCAAGCTTGAATTCCTCGGACGAAAAA 711
DB 334 CTACTTGCTGCTCTCTCTTGTGTATAMKMKWCTTGAATCCKTKVAGARARAAC 275
QY 712 GCAGATCCCTCTCTCTCTCTGCGAGTTGACCTTCAACATTCCTGTGTAATCC 771
DB 274 AYTGTATKTTTTCYACWGRCTTAAGGTAATTCRTTAATTTTWTGTTTAATT 215
QY 772 AGGTCAAGCCCTATTTGCGGCGCTGCTGATCTTCTGAGTCTGTA 818
DB 214 TAATTCGCATATGTTATATGATTTTATTTATGTCGCGCTA 168

RESULT 10
BX425603
LOCUS BX425603 885 bp mRNA linear EST 15-MAY-2003
DEFINITION CX080222A07.3-PRIME, mRNA sequence.
ACCESSION BX425603
VERSION BX425603.1 GI:30770486
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 885)
L4, W.B., Gruber, C., Tesse, J. and Polyes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)

Contact: Genoscope
Genoscope - Centre National de Sequencage

Email: seqref@genoscope.cns.fr; Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of Invitrogen. Contact : Feng Liang Email : fliang@lifetech.com URL : <http://fulllength.invitrogen.com/> Invitrogen Corporation 1600 Faraday Avenue Genoscope, sequence ID : CL0BB022ZA07FP1.

Source

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CR00222A07"
/tissue_type="NEUROBLASTOMA"
/clone_lib="Homo sapiens NEUROBLASTOMA"
/notes="vector: pCMVSORT 6; 1st strand cDNA was primed with a NotI-Oligo(dT) primer. Five prime end enriched double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSORT 6 vector. library was not normalized."

```

Query Match	3.1%;	Score 39.4;	DB 13;	Length 885;
Best Local Similarity	1.5%;	Pred. No. 11;		
Matches	7;	Conservative 147;	Mismatches 323;	Indels 0;
				Gaps 0;

[illegible][illegible]

Ephydrioidae, Drosophilidae; Drosophila.
1 (bases 1 to 995)
Genoscope.
Direct Submission
Genoscope - Centre National de Séquençage :
Submitted (02-JUN-1999) Genoscope / E-mail : secret@genoscope.cns.fr
Submitted (02-JUN-1999) Genoscope / E-mail : secret@genoscope.cns.fr

Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP).

The BDPG is constructing a physical map of the *Drosophila* melanogaster genome using these BACs. For further information please see <http://www.fruitfly.org>. The BDPG *Drosophila* melanogaster BAC library was prepared by Kazutoyo Oseguwa and Aaron Memmoser in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPc1-9 and was constructed by partial EcoRI digestion of *Drosophila* DNA provided by the BDPG from the isogenic strain Y¹; on bp, the same strain used for the BDPG's pl and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

Source

```
1..995
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACR12G04"
/clone_1lb="RPCT-98"
/notes="Tend : Tcr3"
```

Query Match

	Query Match	Similarity	Score	DB	Length	95%	
Best Local	55	3.1%	39.4	29	995		
Matches	55	Conservative	43	Mismatches	65	Indels	0
					Gaps	0	
Qy	1016	CCGCATCAACCTTTCGCGAGCAATCGATAGACTGAAGACTTAACTTTGCTCATTTG	1075				
Db	872	CAGTMTTTCGCGACCGGTSGTGTGMAAAGSASATATGVTGTCGCTTTMTTATTT	813				
Qy	1076	CCGTTGCGATTACAGTGTGGCGCATCTTTTGGCGGTCCAGCGACCACTTTGTGAGCGTTG	1135				
Db	812	CSKTTSSBTWAKASCRMTTTSGSWASMTGCTGKKKGAGSAGAAAGGGCGAKKSYTB	753				
Qy	1136	GGCGTCGACACATCGTTTGTGTCGACTCGTGAATCTTTTCTTA	1178				
Db	752	SSGBTGKTKYSTARRKKTGGTGGSGTGDATMTCTTTGTGTT	710				

RESULT 12	BF341287	859 bp	mRNA	1 linear	EST 22-NOV-2000
LOCUS	60201332221 NCICGAP-Bm64 Homo sapiens		CDNA clone	IMAGE:419244	
DEFINITION	5', mRNA sequence.				
ACCESSION	BF341287				
VERSION	BF341287				
KEYWORDS	BF341287.1	GI:11287778			
SOURCE	EST.				
ORGANISM	Homo sapiens (human)				
	Homo sapiens				

REFERENCE 1 (bases 1 to 859)
AUTHORS NIH-MGC <http://mgc.ncl.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC).
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Stransberg, Ph.D.

Tissue Procurement: David N. Louis, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LINL at:

```

/mol_type="genomic DNA
/db_xref="taxon:7227"
/clone="BACR19D16"
/clone_lib="RPCT-98"
/note="end : TERT"

```

Db 638 SSSSSSTSSSTSTSTSKSSSGSSSSSSSSTTSKSTSAAGSGSWSAAGCGSGGTSTSS 697

QY 355 CCTGCTGTGGGCGACAGCCCTCGGGCGCGCATGCGCGTCAACCATTCGCTGGAAATT 414

[illegible]

475 TTTCATATCGGTTTCCGCTTACGACGAAGCCTACGACGTCGCGGCCAGGCCCG 534

QY 535 AGGCTGTCGGCGCTGGCGAC 554

RESULT 14

DEFINITION	
BX61080 Homo sapiens PLACENTA COT 25-NORMALIZED clone CSOD1079YX16 3-PRIME, mRNA sequence.	Homo sapiens cDNA
BX61080	
BX361080.1 GI:30374504	
EST.	
Homo sapiens (human)	

REFERENCE 1 (bases 1 to 1201)
AUTHORS Li, W.B., Gruber, C., Jesssee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage

Library was constructed by Life Technologies, a division of Invitrogen. Contact : Feng Liang Email : liang@lifetech.com URL : <http://fulllength.invitrogen.com/> Invitrogen Corporation 1600 Faraday Avenue Genoscope sequence ID : CS01079DGC8NP1.

```

SOURCE
1. 1201
/organism="Homo sapiens"
/mol_type="mrna"
/db_xref="txon:9606"
/cidone="CS0D1079YN16"

```

Query Match	ORIGIN
3.0%; Score 38.6; DB 13; Length 1201;	<p> digested with Not I and cloned into the Not I and Xba I sites of the pCMVSPORT 6 vector. Library was normalized. </p>

Matches	34;	Conservative	147;	Mismatches	142;	Indels	0;	Gaps	0;
---------	-----	--------------	------	------------	------	--------	----	------	----

Qy	119	ATTATTCACAGCTGAGAGTGTGGCATGCAAGGAGAGCTGAAACAGATGATTAAGT	178
Db	825	MDDMMMMMMKNNKKKKNDKKMMKKAKKKKKKKDKKKMDKKKKKKKKKAMAKKK	884
Qy	179	TATGGGCGCTACGAATAGCGGAGGTCTAAAAAGCTCCCTTGCTGCAGGTTGGGCGATG	238
Db	885	KKKKKDAKKKKMAKKKKKAAAAAKKKKKKKKKKKMMNNKKKKKKKKKKKKDDMMNN	944
Qy	239	TACCGATGTGATTCGCTTGTCGTCTGTGTTATTCATACAGGCTACGAATGCTGACA	298
Db	945	KKKKKKKKTKKKTTTTTTTTHTTTTTHMMHTTTTTTTTTTTTHNNKKKKKKKKKK	1004
Qy	299	GCCCCACTGTTTTCCGGGCTGATTTTTCGGGGCTCACCGAAATGTGTGTCATGGCCCTC	358
Db	1005	KNDNNMMMMKKKKKKNNKKMMKKKKKKKKNNKKKKNNKKNNKKNNKKNNKKNNKKNN	1064
Qy	359	GTTTGGGCGGACGAGCCCTCTGGGCGCATCGGCTCACACATCTGTGTAACCTTCGCG	418
Db	1065	KKKKKKCCCBKACCCCKKKYCCBKAKKKYCCYKCKKKKKKKKKYCCBCCCKKKCBYCC	1124
Qy	419	CACGATTCATAGCGTTTCATT	441
Db	1125	CCCKKKKKBCKKKKKCCCKKK	1147

RESULT	15
LOCUS	CB657749
DEFINITION	OSJUNEc13F02.f OSJUNEC Oryza sativa (japonica cultivar-group) cDNA
ACCESSION	CB657749
SOURCE	VERSION
KEYWORDS	GI:29661474
ORGANISM	Est.
REFERENCE	Oryza sativa (japonica cultivar-group)
AUTHORS	Oryza sativa (japonica cultivar-group) Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatocoidae; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrenfriedae; Oryzaceae; Oryza. 1 (bases 1 to 835) Jantaaniriyarat,C., Lu,G., Gowd,M., Hatfield,J., Zhou,B., Marur,E., Kudrna,D., Dean,R., Soderlund,C., Ming,R. and Wang,G. large-scale identification of ESTs involved in the interaction between rice and Magnaporthe grisea Unpublished (2003) Contact: Rod Wing Arizona Genomics Institute University of Arizona Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ 85721-0088, USA Tel: 520 626 3967 Fax: 520 621 9288 Email: http://genome.arizona.edu PCR Primers FORWARD: gta aaa cga cgg cca gcg BACKWARD: gga aac agc tat gag cat g Plate: 13 row: F column: 02 Seq primer: gta aaa cga cgg cca gcg. Location/Qualifiers
JOURNAL COMMENT	
TITLE	
FEATURES	
source	.gb35

ORIGIN

```

Query Match: 3.0%; Score 38; DB 14; Length 835;
Best Local Similarity 51.1%; Pred. No. 28;
Matches 89; Conservative 0; Mismatches 85; Indels 0; Gaps 0;

QY 352 CGCGCTCGTTGGGCGCAGAGCCCGCTGGGCGCATATGGCTTACACCAATGTGTGAA 411
Db 79 CGACCGCGCTCCGCGTGGCGCGCTCGTGGGTCTTTCGCGCTGGCGCGAGATGCCCTC 138

QY 412 CTTCGCGCAGATTTCTATGCGTTTATTCCGCTGATGATGATCAAAAACCCCATTCG 471
Db 139 ATCCGCCCGGGGTCTTGCACCGGAATGGCGACCCCGACCTCGCGCGTGGAGACCGCTGCT 198

QY 472 CGCTTTCTATTGGATTTCGCGCTTATCGAGGAAGCTTACGCACTCATCTCGGCG 525
Db 199 CTCCGCGCTACGGCGCTTCGTCGCGCAGGAGCATGGCTTGGCCACCTCTCGGCG 252

Search completed: April 17, 2004, 20:01:40
Job time : 4123.78 secs

```

```

1. .835
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nippohbare"
/db_xref="taxon:3947"
/clone="OSJNEC13F02"
/tissue_type="leaf"
/dev_stage="3 week"
/lab_host="MDH08"
/clone_idb="OSJNEC"
/notes="Vector: pBluescript II KS +; Site_1: EcoRI; Site_2:
XhoI; 6 hrs after inoculation with Rice Blast (C9240-1)"

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 19, 2004, 10:57:18 ; Search time 5387 Seconds

(without alignments)
10226.291 Million cell updates/sec

Title: US-10-608-504-1

Perfect score: 1271

Sequence: 1 ggcgcataacgaagatctag.....agccttaagcagacagcgcct 1271

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 3470272 seqs, 2167151695 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : GenEmbl.*
1: gb_da.*
2: gb_hlg.*
3: gb_in.*
4: gb_cm.*
5: gb_ov.*
6: gb_ph.*
7: gb_pl.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vl.*
15: em_da.*
16: em_fun.*
17: em_hum.*
18: em_in.*
19: em_mu.*
20: em_om.*
21: em_or.*
22: em_ov.*
23: em_pat.*
24: em_ph.*
25: em_pl.*
26: em_ro.*
27: em_sts.*
28: em_un.*
29: em_vl.*
30: em_hlg_hum.*
31: em_hlg_inv.*
32: em_hlg_other.*
33: em_hlg_mus.*
34: em_hlg_pin.*
35: em_hlg_rtd.*
36: em_hlg_mam.*
37: em_hlg_vtl.*
38: em_sy.*
39: em_hgo_hum.*
40: em_hgo_mus.*
41: em_hgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1271	100.0	1271	6 AR391953	AR391953 Sequence
2	1271	100.0	1271	6 AR391956	AR391956 Sequence
3	1271	100.0	1271	6 AX137709	AX137709 Sequence
4	1271	100.0	1271	6 AX137714	AX137714 Sequence
5	1271	100.0	1271	6 BD014991	BD014991 Nucleotide
6	1271	100.0	1271	6 BD014994	BD014994 Nucleotide
7	1271	100.0	340000	6 AP005274	AP005274 Coryneb
8	1271	100.0	349980	6 AX120085	AX120085 Sequence
9	1260	99.1	2105	1 AF454053	AF454053 Coryneb
10	753	59.2	753	6 AR391954	AR391954 Sequence
11	753	59.2	753	6 AX120373	AX120373 Sequence
12	753	59.2	753	6 AX137710	AX137710 Sequence
13	753	59.2	753	6 BD014992	BD014992 Nucleotide
14	753	59.2	753	6 BD162490	BD162490 Novel pol
15	365.6	28.8	397	6 AX137083	AX137083 Sequence
16	365.6	28.8	397	6 AX137085	AX137085 Sequence
17	324	25.5	324	6 AR391955	AR391955 Sequence
18	324	25.5	324	6 AX120374	AX120374 Sequence
19	324	25.5	324	6 AX137712	AX137712 Sequence
20	324	25.5	324	6 BD014993	BD014993 Nucleotide
21	324	25.5	324	6 BD162491	BD162491 Novel pol
22	185	14.6	715	6 AX137077	AX137077 Sequence
23	185	14.6	715	6 BD013995	BD013995 Novel nuc
24	149.6	11.8	87340	1 AP005224	AP005224 Coryneb
25	132.8	10.4	53793	1 BX248361	BX248361 Coryneb
26	127.4	10.0	711	6 AX123492	AX123492 Sequence
27	127.4	10.0	711	6 BD165609	BD165609 Novel pol
28	127.4	10.0	309400	6 AX127153	AX127153 Sequence
29	127.4	10.0	325651	1 AP005283	AP005283 Coryneb
30	59.8	4.7	349926	1 BX571660	BX571660 Wolinella
31	58.6	4.6	756	6 AR386287	AR386287 Sequence
32	57.2	4.5	9888	1 AE011050	AE011050 Methanosa
33	54.4	4.3	856	9 HSA325448	AJ325448 Homo sapi
34	52.4	4.1	7218	6 H64994	166494 Sequence
35	51.8	4.1	11025	1 AE013585	AE013585 Methanosa
36	51.2	4.0	125020	9 AF429315	AF429315 Homo sapi
37	51	4.0	308015	1 AE016783	AE016783 pseudom
38	48.2	3.8	2000	6 AX655393	AX655393 Sequence
39	48	3.8	10732	6 E32986	E32986 Gene encodi
40	46.6	3.7	301838	1 AE017209	AE017209 Geobacter
41	46.6	3.7	3619	1 ECHA10307	AJ410307 Erwinia c
42	46.6	3.7	10733	1 AE013695	AE013695 Yersinia
43	46.6	3.7	220050	1 AJ414156	AJ414156 Yersinia
44	46	3.6	215050	1 AL646084	AL646084 Ralstonia
45	45.8	3.6	20342	1 AE008888	AE008888 Salmonell

ALIGNMENTS

RESULT 1
LOCUS AR391953 1271 bp DNA linear PAT 18-DEC-2003
DEFINITION Sequence 1 from patent US 6613545.
ACCESSION AR391953
VERSION AR391953.1 GI:40115724
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1271)
AUTHORS Kemmerich,N., Sahm,H., Eggeling,L. and Pfeifferle,W.
TITLE Nucleotide sequences coding for the export of branched chain amino acids, process for the isolation thereof and use thereof
JOURNAL Patent: US 6613545-A 1 02-SEP-2003;

FEATURES

source

Location/Qualifiers
1..1271
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN

Query Match 100.0%; Score 1271; DB 6; Length 1271;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1271; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 GGGGATGAATGAGATCTAGCTTCAATATGACCAATAGCCTTGTAGGTGGCAAC 60
DB 1 GGGGATGAATGAGATCTAGCTTCAATATGACCAATAGCCTTGTAGGTGGCAAC 60
QY 61 TGGCAACAAAATCTACCCGGCAATGTGTATGATTTGATGCAAAAACGCAAGAT 120
DB 61 TGGCAACAAAATCTACCCGGCAATGTGTATGATTTGATGCAAAAACGCAAGAT 120
QY 121 TCATTCAGGCTGAGAGGTGTGCGCATCCAGGAGCCCTGGAAACAGATGATAAGTTA 180
DB 121 TCATTCAGGCTGAGAGGTGTGCGCATCCAGGAGCCCTGGAAACAGATGATAAGTTA 180
QY 181 TCGGCGCTACGAATTCGCGCAAGGTCTAATAAATCTCCCTGCTGCAAGTTGGCATGTA 240
DB 181 TCGGCGCTACGAATTCGCGCAAGGTCTAATAAATCTCCCTGCTGCAAGTTGGCATGTA 240
QY 241 CCCGATGGATTTGGGTTTGGTCTCTTGTATTCATATAGGCTACGAATGTTGGGAGC 300
DB 241 CCCGATGGATTTGGGTTTGGTCTCTTGTATTCATATAGGCTACGAATGTTGGGAGC 300
QY 301 CCCAGTGTTCCTGGGCTGATTTTTCGCGGCTCCACCGAAATGCTGTATTCGCTCGT 360
DB 301 CCCAGTGTTCCTGGGCTGATTTTTCGCGGCTCCACCGAAATGCTGTATTCGCTCGT 360
QY 361 TGTGGGCGGAGGCCCCCTGGGCGCATTCGCTCACCATTTGCTGTAATTCGCGCA 420
DB 361 TGTGGGCGGAGGCCCCCTGGGCGCATTCGCTCACCATTTGCTGTAATTCGCGCA 420
QY 421 CGTATTCATGCGTTTTCATTCGCGCTGATGTGTCAAAAACCCCATTTGCCGTTTCTA 480
DB 421 CGTATTCATGCGTTTTCATTCGCGCTGATGTGTCAAAAACCCCATTTGCCGTTTCTA 480
QY 481 TTCGTTTTTCGCGCTTATGACGAAGCTTACGAGTCACTGCGGCGCAGGCGCGAGCTG 540
DB 481 TTCGTTTTTCGCGCTTATGACGAAGCTTACGAGTCACTGCGGCGCAGGCGCGAGCTG 540
QY 541 GTGGGCTGGGAGATTCATCAATGCAATAGGTTTCACTCTGCTGATTTGGGCGG 600
DB 541 GTGGGCTGGGAGATTCATCAATGCAATAGGTTTCACTCTGCTGATTTGGGCGG 600
QY 601 TCTCACCGAGTGGCGATGCGAGAGTGTATCTTTTGAATTAAGGCGCTCGAATTCG 660
DB 601 TCTCACCGAGTGGCGATGCGAGAGTGTATCTTTTGAATTAAGGCGCTCGAATTCG 660
QY 661 CCTTGGCTCTCTTGTGACGCTGACTTGAATTCCTGCGGAGCAAGAAAACAGATCC 720
DB 661 CCTTGGCTCTCTTGTGACGCTGACTTGAATTCCTGCGGAGCAAGAAAACAGATCC 720
QY 721 TTTCTGCTGCTGCAAGGTTTGAAGTTCAATTCCTTGTGATTTCCAGATCAAGC 780
DB 721 TTTCTGCTGCTGCAAGGTTTGAAGTTCAATTCCTTGTGATTTCCAGATCAAGC 780
QY 781 CCTATTTGGGCGGCTGATCTTCTTGGGTCTGTTGACATCCGGATTTCTTTGGG 840
DB 781 CCTATTTGGGCGGCTGATCTTCTTGGGTCTGTTGACATCCGGATTTCTTTGGG 840
QY 841 AAGGCTGCTAATGACAACTGATTTCTCTGATTTCTCTTGTGCTGCAATGATGCA 900
DB 841 AAGGCTGCTAATGACAACTGATTTCTCTGATTTCTCTTGTGCTGCAATGATGCA 900
QY 901 GTATTTACTTTTGGGCTCCGGGCGGTTCCGTTTAACTTAAGCCCTTAAGTGAATCA 960
DB 901 GTATTTACTTTTGGGCTCCGGGCGGTTCCGTTTAACTTAAGCCCTTAAGTGAATCA 960

```

```

QY 961 CAATTTTGGGCAAAATGCGATGATGATGCGACAGAGAAATCTTCCATTTTGAACCGCA 1020
DB 961 CAATTTTGGGCAAAATGCGATGATGATGCGACAGAGAAATCTTCCATTTTGAACCGCA 1020
QY 1021 TCAACGTTTGGAGCAATGCGATGATGATGCGACAGAGAAATCTTCCATTTTGAACCGCA 1080
DB 1021 TCAACGTTTGGAGCAATGCGATGATGATGCGACAGAGAAATCTTCCATTTTGAACCGCA 1080
QY 1081 GCGATTCAGAGTGTGCGCATCTTCTTGGCGTCAAGCAGCCTTGTGACGTTGGCGCT 1140
DB 1081 GCGATTCAGAGTGTGCGCATCTTCTTGGCGTCAAGCAGCCTTGTGACGTTGGCGCT 1140
QY 1141 GGCACCATGTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1200
DB 1141 GGCACCATGTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1200
QY 1201 CCGCATGCTCTCAATTTGAAGGAGATGCGGATTTTGAAGAACTTGAAGAAAGCTTAAG 1260
DB 1201 CCGCATGCTCTCAATTTGAAGGAGATGCGGATTTTGAAGAACTTGAAGAAAGCTTAAG 1260
QY 1261 CAGACAGCGCT 1271
DB 1261 CAGACAGCGCT 1271

```

RESULT 2

AR391956 1271 bp DNA linear PAT 18-DEC-2003
LOCUS AR391956
DEFINITION Sequence 6 from patent US 6613545.
ACCESSION AR391956
VERSION AR391956.1 GI:40115727
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1271)
AUTHORS Kennertknecht,N., Sahm,H., Eggeling,L. and Pfeifferle,W.
TITLE Nucleotide sequences coding for the export of branched chain amino acids, process for the isolation thereof and use thereof
JOURNAL Patent: US 6613545-A 6 02-SEP-2003;
FEATURES
source
1..1271
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN

Query Match 100.0%; Score 1271; DB 6; Length 1271;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1271; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 GGGGATGAATGAGATCTAGCTTCAATATGACCAATAGCCTTGTAGGTGGCAAC 60
DB 1 GGGGATGAATGAGATCTAGCTTCAATATGACCAATAGCCTTGTAGGTGGCAAC 60
QY 61 TGGCAACAAAATCTACCCGGCAATGTGTATGATTTGATGCAAAAACGCAAGAT 120
DB 61 TGGCAACAAAATCTACCCGGCAATGTGTATGATTTGATGCAAAAACGCAAGAT 120
QY 121 TCATTCAGGCTGAGAGGTGTGCGCATCCAGGAGCCCTGGAAACAGATGATAAGTTA 180
DB 121 TCATTCAGGCTGAGAGGTGTGCGCATCCAGGAGCCCTGGAAACAGATGATAAGTTA 180
QY 181 TCGGCGCTACGAATTCGCGCAAGGTCTAATAAATCTCCCTGCTGCAAGTTGGCATGTA 240
DB 181 TCGGCGCTACGAATTCGCGCAAGGTCTAATAAATCTCCCTGCTGCAAGTTGGCATGTA 240
QY 241 CCCGATGGATTTGGGTTTGGTCTCTTGTATTCATATAGGCTACGAATGTTGGGAGC 300
DB 241 CCCGATGGATTTGGGTTTGGTCTCTTGTATTCATATAGGCTACGAATGTTGGGAGC 300
QY 301 CCCAGTGTTCCTGGGCTGATTTTTCGCGGCTCCACCGAAATGCTGTATTCGCTCGT 360
DB 301 CCCAGTGTTCCTGGGCTGATTTTTCGCGGCTCCACCGAAATGCTGTATTCGCTCGT 360

```

[illegible]

KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source

Corynebacterium glutamicum
Corynebacterium glutamicum
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacteriaceae; Corynebacteriaceae; Corynebacterium.
1
Kamerknecht, N., Eggeling, L., Sahm, H. and Pfeifferle, W.
Nucleotide sequences coding for branched-chain amino acids export
proteins, method for isolating them and their use
Patent: EP 1096010-A 1 02-MAY-2001;
Degussa AG (DE) / FORSCHUNGSZENTRUM JUELICH GMBH (DE)
Local/on/Qualifiers

1. 1271
/organism="Corynebacterium glutamicum"
/mol_type="unassigned DNA"
/db_xref="taxon:1718"
/note="ATCC14752"
101..856
/note="unnamed protein product; brnf"
/codon_start=1
/transl_table=1
/protein_id="CAC39923.1"
/db_xref="GI:14273887"
/db_xref="REMRMBL:CAC39923"
/translation="MKTQETHSLSVPSKRALPPDDKGRVREINOGKLSIAGI
GMPIGIAGFLVQYETEMAPLSEGLIFAGSTEMVLALVGAAPLGIATLII
YENFHVAYASFLPHVAVKNPIARYSVFALIDEVATYAPAPMSMRLLSWIARF
SYWVFGGLTGVAILNELPPEIKGLEPFLCSFLVLTLDSCRTKQIPSLILLAGSFTT
ALVVPQALFPAALILFLGLITIRYFELGRAAK"
853..1179
/note="unnamed protein product; brnf"
/codon_start=1
/transl_table=1
/protein_id="CAC39924.1"
/db_xref="GI:14273888"
/db_xref="REMRMBL:CAC39924"
/translation="MTDPSGILLVAVCAVITPALRAVPLILKPRESOFGKMA
WMPGILAILLASVFRSNALDKLTFTGLIAVAITVAHLGLGRITLSVAGATVFA
GLVNLF"

ORIGIN

Query Match 100.0%; Score 1271; DB 6; Length 1271;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1271; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 GCGGATCAATGGAATCTAGCTTCATATATATGCAACATAGCTCTAGGTGGGCAAC 60
1 GCGGATCAATGGAATCTAGCTTCATATATGCAACATAGCTCTAGGTGGGCAAC 60
QY TGGCAACAAACTACCCGGCAATGTGTGATGATGATTTGTAGTGTGCAAAAAACGAAGAT 120
61 TGGCAACAAACTACCCGGCAATGTGTGATGATGATTTGTAGTGTGCAAAAAACGAAGAT 120
Db TGGCAACAAACTACCCGGCAATGTGTGATGATGATTTGTAGTGTGCAAAAAACGAAGAT 120
61 TGGCAACAAACTACCCGGCAATGTGTGATGATGATTTGTAGTGTGCAAAAAACGAAGAT 120
QY TCGATTCAAGCTGTGAAGTGTGCCATCCAGCAAGCCCTGTGAACCAAGATGAAGAAGTTA 180
121 TCGATTCAAGCTGTGAAGTGTGCCATCCAGCAAGCCCTGTGAACCAAGATGAAGAAGTTA 180
121 TCGATTCAAGCTGTGAAGTGTGCCATCCAGCAAGCCCTGTGAACCAAGATGAAGAAGTTA 180
Db TCGATTCAAGCTGTGAAGTGTGCCATCCAGCAAGCCCTGTGAACCAAGATGAAGAAGTTA 180
121 TCGATTCAAGCTGTGAAGTGTGCCATCCAGCAAGCCCTGTGAACCAAGATGAAGAAGTTA 180
QY TCGGGGCTACGAATCGCGCAAGGCTTAACAACTCCCTTGTGTGCAAGTTTGGGCGATGTA 240
181 TCGGGGCTACGAATCGCGCAAGGCTTAACAACTCCCTTGTGTGCAAGTTTGGGCGATGTA 240
181 TCGGGGCTACGAATCGCGCAAGGCTTAACAACTCCCTTGTGTGCAAGTTTGGGCGATGTA 240
Db TCGGGGCTACGAATCGCGCAAGGCTTAACAACTCCCTTGTGTGCAAGTTTGGGCGATGTA 240
181 TCGGGGCTACGAATCGCGCAAGGCTTAACAACTCCCTTGTGTGCAAGTTTGGGCGATGTA 240
QY CCCGATTGGTATTTGGCTTCTCTTGGTTATTCAATAGCGCTACGAATGGTGGGCAAGC 300
241 CCCGATTGGTATTTGGCTTCTCTTGGTTATTCAATAGCGCTACGAATGGTGGGCAAGC 300
241 CCCGATTGGTATTTGGCTTCTCTTGGTTATTCAATAGCGCTACGAATGGTGGGCAAGC 300
Db CCCGATTGGTATTTGGCTTCTCTTGGTTATTCAATAGCGCTACGAATGGTGGGCAAGC 300
241 CCCGATTGGTATTTGGCTTCTCTTGGTTATTCAATAGCGCTACGAATGGTGGGCAAGC 300
QY CCCGATTGGTATTTGGCTTCTCTTGGTTATTCAATAGCGCTACGAATGGTGGGCAAGC 300
301 CCCGATTGGTATTTGGCTTCTCTTGGTTATTCAATAGCGCTACGAATGGTGGGCAAGC 300
301 CCCGATTGGTATTTGGCTTCTCTTGGTTATTCAATAGCGCTACGAATGGTGGGCAAGC 300
Db CCCGATTGGTATTTGGCTTCTCTTGGTTATTCAATAGCGCTACGAATGGTGGGCAAGC 300
301 CCCGATTGGTATTTGGCTTCTCTTGGTTATTCAATAGCGCTACGAATGGTGGGCAAGC 300
QY TGTGGGGCGAGCGCCCTGTGGGCGCCATGCGCTACCAACCAATTGCTGTGAACCTTCCGCA 420
361 TGTGGGGCGAGCGCCCTGTGGGCGCCATGCGCTACCAACCAATTGCTGTGAACCTTCCGCA 420
361 TGTGGGGCGAGCGCCCTGTGGGCGCCATGCGCTACCAACCAATTGCTGTGAACCTTCCGCA 420
Db TGTGGGGCGAGCGCCCTGTGGGCGCCATGCGCTACCAACCAATTGCTGTGAACCTTCCGCA 420
361 TGTGGGGCGAGCGCCCTGTGGGCGCCATGCGCTACCAACCAATTGCTGTGAACCTTCCGCA 420
361 TGTGGGGCGAGCGCCCTGTGGGCGCCATGCGCTACCAACCAATTGCTGTGAACCTTCCGCA 420

QY	421	CGAATCTAAGCGTTTCAATTCGCGCTGAGATGAGTCAAAAACCCATTCGCGTTTCA	480
Db	421	CGAATTCATGCGTTTTCATTCGCGCTGAGATGAGTCAAAAACCCATTCGCGTTTCA	480
QY	481	TTGCGTTTTCGCGCTTATCGACGAAGCTTACGAGTCACTTCGCGCCAGCTCGAGCTTG	540
Db	481	TTGCGTTTTCGCGCTTATCGACGAAGCTTACGAGTCACTTCGCGCCAGCTCGAGCTTG	540
QY	541	GTGCGCGTGGGAGATTATCTCAAGCAATAAGCGTTCACTCCACTGAGGGATTTGCGCGG	600
Db	541	GTGCGCGTGGGAGATTATCTCAAGCAATAAGCGTTCACTCCACTGAGGGATTTGCGCGG	600
QY	601	TCTCACCGGAGTGGCGCATGCGAGAGTGAATTCCTTTGAAATTAAAGGCGCTCGAGTTGCG	660
Db	601	TCTCACCGGAGTGGCGCATGCGAGAGTGAATTCCTTTGAAATTAAAGGCGCTCGAGTTGCG	660
QY	661	CCTTTGCTCTCTCTTTGTCAAGCTGACTTTGGATTTCTCTGCGGAACGAAAAAGCAGATCCC	720
Db	661	CCTTTGCTCTCTCTTTGTCAAGCTGACTTTGGATTTCTCTGCGGAACGAAAAAGCAGATCCC	720
QY	721	TTCTGTGCTGTGAGAGTTTGAAGCTTCAACCATGCTCTTGAGTAATTCAGGTCAAGC	780
Db	721	TTCTGTGCTGTGAGAGTTTGAAGCTTCAACCATGCTCTTGAGTAATTCAGGTCAAGC	780
QY	781	CCCTATTGCGGCGCTGATCTTCTGGGTCCTGTGAACATCCGGTACTTCTTTGGG	840
Db	781	CCCTATTGCGGCGCTGATCTTCTGGGTCCTGTGAACATCCGGTACTTCTTTGGG	840
QY	841	AAAGCGCTAAATGACAACTGATTTCTCCGTAATTCCTGTTGTGCGAGTATGTGCA	900
Db	841	AAAGCGCTAAATGACAACTGATTTCTCCGTAATTCCTGTTGTGCGAGTATGTGCA	900
QY	901	GTCATTTACTTTTGGCGCTCCGCGCGGTTCCGTTCTTAATCTTAAAGCCCTACGTAAACA	960
Db	901	GTCATTTACTTTTGGCGCTCCGCGCGGTTCCGTTCTTAATCTTAAAGCCCTACGTAAACA	960
QY	961	CAATTTGGGGCAAAATGGCGATGTGATGTCCAGAGAAATCCCTTGCAATTTGACCGCA	1020
Db	961	CAATTTGGGGCAAAATGGCGATGTGATGTCCAGAGAAATCCCTTGCAATTTGACCGCA	1020
QY	1021	TCAAAGTTGCGAGCATGCGATAGATGTGAAGACTTAAACCTTTGATCTCATTTCCGTT	1080
Db	1021	TCAAAGTTGCGAGCATGCGATAGATGTGAAGACTTAAACCTTTGATCTCATTTCCGTT	1080
QY	1081	GCGATTAAGATGTGTGGCGCATTTCTTTGGCGGTCAAGCAGCTTGTAGAGTTGGCGCT	1140
Db	1081	GCGATTAAGATGTGTGGCGCATTTCTTTGGCGGTCAAGCAGCTTGTAGAGTTGGCGCT	1140
QY	1141	GCGACCAATCGTTTGTGTGGACTGTGTGAATCTTTCTTAAACTGATTAATACAAAAAT	1200
Db	1141	GCGACCAATCGTTTGTGTGGACTGTGTGAATCTTTCTTAAACTGATTAATTAACAAAAAT	1200
QY	1201	CCGCAATGCCCTCAATTTGAAGGGAGTGCAGATTTTAAAGAACTTAAGAAAGGCTTAAAG	1260
Db	1201	CCGCAATGCCCTCAATTTGAAGGGAGTGCAGATTTTAAAGAACTTAAGAAAGGCTTAAAG	1260
QY	1261	CAGACAGCGCT 1271	
Db	1261	CAGACAGCGCT 1271	

RESULT 4				
AX137714				
LOCUS	AX137714	1271 bp	DNA	linear
DEFINITION	Sequence 6 from Patent EP1096010.			PAT 30-MAY-2001
ACCESSION	AX137714			
VERSION	AX137714.1	GI:14273893		
KEYWORDS	.			
SOURCE	Corynebacterium glutamicum			
ORGANISM	Corynebacterium glutamicum			
	Bacteria; Actinobacterii; Actinobacteridae; Actinomycetales;			
	Corynebacteriaceae; Corynebacteriinae; Corynebacterium.			
REFERENCE	1			

ORIGIN	Query Match	Best Local Similarity	Matches 1271, Conservative	Score 1271; DB 6; Length 1271; 100.0%; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0
QY	1	GCGGATCAATGGGATCTAGCTCATATATATGCAACATAGCCTAGTTGAGTGCCGAAC	60	
Db	1	GCGGATCAATGGGATCTAGCTCATATATGCAACATAGCCTAGTTGAGTGCCGAAC	60	
QY	61	TGGCAACAAACCTACCCGGCAATTGTGTGATGATTTGTGTGCAAAAAACGAGAAGAT	120	
Db	61	TGGCAACAAACCTACCCGGCAATTGTGTGATGATTTGTGTGCAAAAAACGAGAAGAT	120	
QY	121	TCATTCAAGCCTGGAGGTGTGCGCCATCCAAAGGACCCCTGGAAACAGATGATAAGTTA	180	
Db	121	TCATTCAAGCCTGGAGGTGTGCGCCATCCAAAGGACCCCTGGAAACAGATGATAAGTTA	180	
QY	181	TCGGGCTAGCAATTCGCGCAAGTCTTAAAAACCTCCCTGTGTGCAAGTTTGGGATGTA	240	
Db	181	TCGGGCTAGCAATTCGCGCAAGTCTTAAAAACCTCCCTGTGTGCAAGTTTGGGATGTA	240	
QY	241	CCGCAATGATATGGGTTTGGCTCTTGGTATTCATATACGGCTACGAATGCTGGGACAC	300	
Db	241	CCGCAATGATATGGGTTTGGCTCTTGGTATTCATATACGGCTACGAATGCTGGGACAC	300	
QY	301	CCGACTGTTTTCCGGCTGATTTTTCGGGGGCTCCACCGAATGCTGGTATGCGCTCGT	360	
Db	301	CCGACTGTTTTCCGGCTGATTTTTCGGGGGCTCCACCGAATGCTGGTATGCGCTCGT	360	
QY	361	TGTGGGCGAGGCGCCCTGGGGCGCCATGCGCTGCACCATGCTGGTGAATTCGCGCA	420	
Db	361	TGTGGGCGAGGCGCCCTGGGGCGCCATGCGCTGCACCATGCTGGTGAATTCGCGCA	420	
QY	421	CGTATTCATGCGTTTCAATTCGCGCTGCATGTGTCAAAAAACCCGATTGCGCGTTTCTA	480	
Db	421	CGTATTCATGCGTTTCAATTCGCGCTGCATGTGTCAAAAAACCCGATTGCGCGTTTCTA	480	
QY	481	TTGGATTTTCGGGCTTATGACGAAGCTACGCAAGTCACTGGGGCGACGCGCAGGCTG	540	

Db 481 TTCGATTTTCGCGCTTATCGACGAAGCCTACGACGACTGCGGCCAGGCCGCGAGGCTG 540
 Qy 541 GTGCGCGCGGAGCTTATCTCAATGCAATAGCGTTTCACTCCACTAGGGGATTCGGCGG 600
 Db 541 GTGCGCGCGGAGCTTATCTCAATGCAATAGCGTTTCACTCCACTAGGGGATTCGGCGG 600
 Qy 601 TCTGACCGGAGTGGCGATGCGAGATTGATCTTTGAAATTAAGGGCCCTCGATTGCG 660
 Db 601 TCTGACCGGAGTGGCGATGCGAGATTGATCTTTGAAATTAAGGGCCCTCGATTGCG 660
 Qy 661 CTTTGTCTCTCTTTGTCAAGCTGACTTTGATTCCTGCGGAAAGCAAAAGCGATCCC 720
 Db 661 CTTTGTCTCTCTTTGTCAAGCTGACTTTGATTCCTGCGGAAAGCAAAAGCGATCCC 720
 Qy 721 TTTCTGTGTCGCGAGGTTTGAAGCTTCAACATTCCTTTGAGTAATTCAGGTCAGGC 780
 Db 721 TTTCTGTGTCGCGAGGTTTGAAGCTTCAACATTCCTTTGAGTAATTCAGGTCAGGC 780
 Qy 781 CCTATTGCGGCGCTGCTGATCTTCTGAGTCTGTTGACCAATCCGATCCTTCTTGGG 840
 Db 781 CCTATTGCGGCGCTGCTGATCTTCTGAGTCTGTTGACCAATCCGATCCTTCTTGGG 840
 Qy 841 AAGGCTCTAAATGACAACTGATTTCTCTGATTCCTGTTGCGAGTAATGCA 900
 Db 841 AAGGCTCTAAATGACAACTGATTTCTCTGATTCCTGTTGCGAGTAATGCA 900
 Qy 901 GTCATTTACCTTTGCGCTCCGGCGGATTCCTTCTTAATCCTTAAGCCCTTACGTAATCA 960
 Db 901 GTCATTTACCTTTGCGCTCCGGCGGATTCCTTCTTAATCCTTAAGCCCTTACGTAATCA 960
 Qy 961 CAATTGTTGGGGAATGGCGATGCGATGCGAGGAAATCCCTTGAATTTGACGCA 1020
 Db 961 CAATTGTTGGGGAATGGCGATGCGATGCGAGGAAATCCCTTGAATTTGACGCA 1020
 Qy 1021 TCAAGCTTTCGAGCAATGCGATGATCTGAAGACTTAACCTTGTCTGATTCGCGCT 1080
 Db 1021 TCAAGCTTTCGAGCAATGCGATGATCTGAAGACTTAACCTTGTCTGATTCGCGCT 1080
 Qy 1081 GCGATTACAGTGTGGCGCATTTCTTGGCGGTGACGCACTTTGTAAGGTTGGCGCT 1140
 Db 1081 GCGATTACAGTGTGGCGCATTTCTTGGCGGTGACGCACTTTGTAAGGTTGGCGCT 1140
 Qy 1141 GGCACCATGTTTGTGTTGGAATGCTGTTTCTAAACTGATTAATTAACAAAT 1200
 Db 1141 GGCACCATGTTTGTGTTGGAATGCTGTTTCTAAACTGATTAATTAACAAAT 1200
 Qy 1201 CCGCATGCGCTCAATTGAAAGGGAATGCGGATTTTAAAGAACTTAAGAAAGGCTTAAG 1260
 Db 1201 CCGCATGCGCTCAATTGAAAGGGAATGCGGATTTTAAAGAACTTAAGAAAGGCTTAAG 1260
 Qy 1261 CAGACAGCGCT 1271
 Db 1261 CAGACAGCGCT 1271
 RESULT 5
 LOCUS BD014991 1271 bp DNA linear PAT 27-AUG-2002
 DEFINITION Nucleotide sequence encoding the discharge of branched amino acid
 method of isolating the same and utilization thereof.
 ACCESSION BD014991 GI:2255798
 VERSION BD014991.1
 KEYWORDS JP 2001169788-A/1.
 SOURCE unclassified
 ORGANISM unclassified
 REFERENCE 1 (bases 1 to 1271)
 AUTHORS Kernerukunehito, N., Sahn, H., Eggering, L. and Pfeifferle, W.
 TITLE Nucleotide sequence encoding the discharge of branched amino acid,
 method of isolating the same and utilization thereof
 JOURNAL Patent: JP 2001169788-A 1 26-JUN-2001;
 DEUSSA HUELS AG FORSCHUNGSZENTRUM JUELICH GMBH
 COMMENT OS Corynebacterium glutamicum ATCC14572

Handwritten signature/initials

PN JP 2001169788-A/1
 PD 26-JUN-2001
 PF 24-OCT-2000 JP 2000324315
 PR 27-OCT-1999 DE 19951708 8
 PI NICOLE KERNERUKUNEHITO, HERMANN SAHN, LOTHAR EGGERING, WALTER PI
 PFEIFFERLE
 PC C12N15/09, C07K14/34, C12N1/21, C12P13/06, C12P13/06, C12P13/08//
 (C12N15/09, C12R1:15), (C12N1/21, C12R1:15), (C12P13/06, C12R1:15),
 PC C12N15/00,
 PC (C12N15/00, C12R1:15)
 CC brnE
 CC brnE
 FH Key
 FT gene
 FT gene
 FEATURES
 source 1..1271
 /organism="unclassified"
 /mol_type="genomic DNA"
 /db_xref="taxon:32644"
 ORIGIN
 Query Match 100.0%; Score 1271; DB 6; Length 1271;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1271; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 GCGCGATCAATGGAATCTAGCTTCAATATATGCAAAATAGCTTGAAGTGGCAAC 60
 Db 1 GCGCGATCAATGGAATCTAGCTTCAATATATGCAAAATAGCTTGAAGTGGCAAC 60
 Qy 61 TGGCAACAAACCTAACCCGGAATGTGTGATGATGTTAGTGTGCAAAACGCAAGAT 120
 Db 61 TGGCAACAAACCTAACCCGGAATGTGTGATGATGTTAGTGTGCAAAACGCAAGAT 120
 Qy 121 TCAATGAGCTGAGAGTGTGCGCATCAAGGAGCCCTGGAACCAAGATTAAGCTTA 180
 Db 121 TCAATGAGCTGAGAGTGTGCGCATCAAGGAGCCCTGGAACCAAGATTAAGCTTA 180
 Qy 181 TGGCGGTAGGAATGGCGGAAGTCTAAAAACCTCCCTTGTGCGAGGTTGGGCAATGA 240
 Db 181 TGGCGGTAGGAATGGCGGAAGTCTAAAAACCTCCCTTGTGCGAGGTTGGGCAATGA 240
 Qy 181 TGGCGGTAGGAATGGCGGAAGTCTAAAAACCTCCCTTGTGCGAGGTTGGGCAATGA 240
 Db 181 TGGCGGTAGGAATGGCGGAAGTCTAAAAACCTCCCTTGTGCGAGGTTGGGCAATGA 240
 Qy 241 CCGGATGATTTGCGTTGCTGCTGTTGATTAATACGCTTGAAGATGCTGAGGCAAG 300
 Db 241 CCGGATGATTTGCGTTGCTGCTGTTGATTAATACGCTTGAAGATGCTGAGGCAAG 300
 Qy 241 CCGGATGATTTGCGTTGCTGCTGTTGATTAATACGCTTGAAGATGCTGAGGCAAG 300
 Db 241 CCGGATGATTTGCGTTGCTGCTGTTGATTAATACGCTTGAAGATGCTGAGGCAAG 300
 Qy 301 CCGCATGTTTCCGGCTGATTTTCCGGGCTCCACCGAAATGCTGCTATGCGCCCTCGT 360
 Db 301 CCGCATGTTTCCGGCTGATTTTCCGGGCTCCACCGAAATGCTGCTATGCGCCCTCGT 360
 Qy 301 CCGCATGTTTCCGGCTGATTTTCCGGGCTCCACCGAAATGCTGCTATGCGCCCTCGT 360
 Db 301 CCGCATGTTTCCGGCTGATTTTCCGGGCTCCACCGAAATGCTGCTATGCGCCCTCGT 360
 Qy 361 TGTGGGCGAGCGCCCTGAGGCGCATGCGGCTCACACCAATGCTGTGTAATTCGCGCA 420
 Db 361 TGTGGGCGAGCGCCCTGAGGCGCATGCGGCTCACACCAATGCTGTGTAATTCGCGCA 420
 Qy 361 TGTGGGCGAGCGCCCTGAGGCGCATGCGGCTCACACCAATGCTGTGTAATTCGCGCA 420
 Db 361 TGTGGGCGAGCGCCCTGAGGCGCATGCGGCTCACACCAATGCTGTGTAATTCGCGCA 420
 Qy 421 CGATTTCTATGCGTTTCAATCCCGCTGCAATGCTGCAAAACCCCATTCGCGTTCTA 480
 Db 421 CGATTTCTATGCGTTTCAATCCCGCTGCAATGCTGCAAAACCCCATTCGCGTTCTA 480
 Qy 421 CGATTTCTATGCGTTTCAATCCCGCTGCAATGCTGCAAAACCCCATTCGCGTTCTA 480
 Db 421 CGATTTCTATGCGTTTCAATCCCGCTGCAATGCTGCAAAACCCCATTCGCGTTCTA 480
 Qy 481 TTTGGTTTGGCGTTTGAAGCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCTT 540
 Db 481 TTTGGTTTGGCGTTTGAAGCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCTT 540
 Qy 481 TTTGGTTTGGCGTTTGAAGCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCTT 540
 Db 481 TTTGGTTTGGCGTTTGAAGCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCTT 540
 Qy 541 GTGCGGTGGCGACTTATCTCAATGCAATAGCGTTTCACTCTGATGAGTATTCGCGG 600
 Db 541 GTGCGGTGGCGACTTATCTCAATGCAATAGCGTTTCACTCTGATGAGTATTCGCGG 600
 Qy 541 GTGCGGTGGCGACTTATCTCAATGCAATAGCGTTTCACTCTGATGAGTATTCGCGG 600
 Db 541 GTGCGGTGGCGACTTATCTCAATGCAATAGCGTTTCACTCTGATGAGTATTCGCGG 600
 Qy 601 TCTGACCGGAGTGGCGATGCGAGATTGATCTTTGAAATTAAGGGCCCTCGATTGCG 660
 Db 601 TCTGACCGGAGTGGCGATGCGAGATTGATCTTTGAAATTAAGGGCCCTCGATTGCG 660
 Qy 661 CTTTGTCTCTCTTTGTCAAGCTGACTTTGATTCCTGCGGAAAGCAAAAGCGATCCC 720
 Db 661 CTTTGTCTCTCTTTGTCAAGCTGACTTTGATTCCTGCGGAAAGCAAAAGCGATCCC 720

QY 901 GTCATTACTTTTGGCGCTCCGGGCGGTTCCGTTCTTAATCTTAAGCCCTACGTGAATCA 960
Db 901 GTCATTACTTTTGGCGCTCCGGGCGGTTCCGTTCTTAATCTTAAGCCCTACGTGAATCA 960
QY 961 CATTTTGGGGGAAAATGGGAGTGTGGATGCCGAGCGGAATCCTTGGCAATTTTGACCGGA 1020
Db 961 CATTTTGGGGGAAAATGGGAGTGTGGATGCCGAGCGGAATCCTTGGCAATTTTGACCGGA 1020
QY 1021 TCAACGTTTGGAGCAATGGGATAGATCTGAAGACTCTTAACCTTGGTCTCATTTGCCGTT 1080
Db 1021 TCAACGTTTGGAGCAATGGGATAGATCTGAAGACTCTTAACCTTGGTCTCATTTGCCGTT 1080
QY 1081 GCGATTACAGTGGGCGGCAATCTTTGGGCGGTGAGAGCACTTTTGAACGTTGGCGCT 1140
Db 1081 GCGATTACAGTGGGCGGCAATCTTTGGGCGGTGAGAGCACTTTTGAACGTTGGCGCT 1140
QY 1141 GGCACATCGTTTGTGTGGACTGTGATCTTTTCTAAACTGCAATTAATACAAAAT 1200
Db 1141 GGCACATCGTTTGTGTGGACTGTGATCTTTTCTAAACTGCAATTAATACAAAAT 1200
QY 1201 CCGCATCCCTCAATTGGAAGGGGATGCGGATTTTGAAGCACTTGAAGAGCTTTAG 1260
Db 1201 CCGCATCCCTCAATTGGAAGGGGATGCGGATTTTGAAGCACTTGAAGAGCTTTAG 1260
QY 1261 CAGACAGCGCT 1271
Db 1261 CAGACAGCGCT 1271

RESULT 7
AP005274 340000 bp DNA linear BCT 08-AUG-2002
LOCUS
DEFINITION Corynebacterium glutamicum ATCC 13032 DNA, complete genome, section 1/10.
ACCESSION AP005274 BA000036
VERSION AP005274.1 GI:21322764
KEYWORDS
SOURCE
ORGANISM Corynebacterium glutamicum ATCC 13032
Corynebacterium glutamicum ATCC 13032
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacteriaceae; Corynebacteriaceae; Corynebacterium.
REFERENCE 1
AUTHORS Nakagawa, S.
TITLE Complete genomic sequence of Corynebacterium glutamicum ATCC 13032
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 340000)
AUTHORS Nakagawa, S.
TITLE Direct Submission
JOURNAL Submitted (24-MAY-2002) Satoshi Nakagawa, Kyowa HAKKO Kogyo Co. Ltd., Tokyo Research Laboratories; 3-6-6, Asahi-machi, Machida, Tokyo 194-8533, Japan (E-mail:snakagawa@xanagen.com, Tel:81-44-829-3031, Fax:81-44-813-1651)
COMMENT This sequence is conducted by collaboration of Kyowa HAKKO Kogyo Co. Ltd. and Kitesato University.
FEATURES
source
1. 340000
/organism="Corynebacterium glutamicum ATCC 13032"
/mol_type="genomic DNA"
/strain="ATCC 13032"
/db_xref="taxon:196627"
/note="ATCC 13032"
1. 1575
/gene="Cg10001"
1. 1575
/gene="Cg10001"
1. 1575
/note="PF00308:Bacterial dnaA protein
TIGR00362:dnaA: chromosomal replication initiator protein
dnaA"
/codon_start=1
/transl_table=11
/product="ATPase involved in DNA replication initiation"
/protein_id="BAB97394.1"
/db_xref="GI:21322765"

/translation="MSQSSSILETRQVAVADLTLSQADSGFDLPITPQRAYLNT
KPIAVGVAVLSTPNAKAVIENDJADLTRLSLRMGSRFLAASVREBOIPEP
PAQOEPKQDPAPVSISSNAPKQYEVGREGASLSDSDEHPAPAPAPAPADPE
PELATPQIPREPAPHPNPNREVSLNPKYTESFVIGPNRPANAAVAAPAPAPAE
PLFISGSGLGKTHLHAGVNAQELQGLIKVSSSEPNIDINSRDROQTEPR
RYRNDILMVDIOFLAGKEGTOEFETFNALQADKOILSSDRPKQTLLEDR
RTREPGCLITDIOCPDETRIALIMKKAQGTGTVDRLELISREFSSIRELGA
IRVSAVSSLIQPIIDKEMALVALRDILPEPDMETTPAVIMEVTAAYEISVDTLRGA
GTRAVARARQALATLCHELTDMSLPKIGDVFSGKDHITVYAPRKIRQENTBERDIT
DEIQQLTQLIKSRGN"
/gene="Cg10002"
complement (1594..1920)
/gene="Cg10002"
/gene="Cg10002"
/codon_start=1
/transl_table=11
/product="Hypothetical protein"
/protein_id="BAB97395.1"
/db_xref="GI:21322766"
/translation="MIVDNFVYLMIDFKVRLSPVCADVSPAYPOFLGSCAQGKFRG
ITFPYTKLTELQVCNYTAVNKVVDNFVSMATIGVSPQNPCTTPKLQKPSDEGSRRA
ELNFIY"
2292..3476
/gene="Cg10003"
2292..3476
/gene="Cg10003"
/EC_number="2.7.7.7"
/note="PF00712:DNA polymerase III beta subunit, N-terminal
domain
PF02767:DNA polymerase III beta subunit, central domain
PF02768:DNA polymerase III beta subunit, C-terminal domain
TIGR00663:dnaA: DNA polymerase III, beta subunit"
/codon_start=1
/transl_table=11
/product="DNA polymerase III beta subunit"
/protein_id="BAB97396.1"
/db_xref="GI:21322767"
/translation="MESQNSFVRVAREDELTVAVAVANLPTKPTQPLVRLMLTAD
BELEAGYDVSTVRLSFAEVSQGRVAVAGKLSEITGSLPNKPVFRIDGSAFV
TCGSSRFELPLIDDYPMPLPGLPAGTGSINAKLFTBEASQVASAGGDDSLPMLTG
SMEIVGNOINLADTRFLALPTFMEANDPELVNKLIPATTLIDNARSLSGGLNDS
IDIAVGTDGVQREGLFGVHTDNRRTTRLDADPNIALPLKEHTAIASVEIAPLV
DAIRVSVIABERNQIVLHFSQVILTAGTAGHARETTPCAFTGKELTIAFPGY
LNDGSLVFTSRAVFGFTPSRPALMITPEPMSANENGITQTDTYFTYIMVRL
PG"
3585..4769
/gene="Cg10004"
3585..4769
/note="TIGR00611:recf: Recf protein"
/codon_start=1
/transl_table=11
/product="Recombinational DNA repair ATPase"
/protein_id="BAB97397.1"
/db_xref="GI:21322768"
/translation="MHTISLELDYRSWPELVKDDPGTTFVIGNGRQKTNVIAIG
YLAHSRHSVSDAPLVRHAHNAEDVSAVAVNGEGLAHILIKHPANQASLNTRKY
RTPRELIGVAKTVLAPEDIALYKBPERRRYDDITATRPRAAGADYDXTKLO
RNALNTATIALRIGTGEBGAALSLDTWGTGARGUAEVAAWAPALNLEGRKIX
EAYTTIAPESRPAAVYKTTIDQGLSQSEFPDAGIIEATLTLELAQRRETERSS
VGPBHDVDLMGDQPAKGFASGETWSFALRLAEFNLISLSDTDLTLLIDVFS
LDAGREKLVGIAQVEQVLTAAVHDLPEMLKKVLVQHTVTVQDNGTGRISLIDV
OP"
4766..5302
/gene="Cg10005"
4766..5302
/gene="Cg10005"
/codon_start=1
/transl_table=11
/product="Hypothetical protein"
/protein_id="BAB97398.1"
/db_xref="GI:21322769"
/translation="MTDPIEQAFERIRAEARNRNSVDPOLKNDAFRPPAPKGVEX

RRKGRASGLDGRQRYVGAESLGSVLNKEIQRGCGKDIAGMGVTSNMBELVGAKIA
OHTREVMIKDKKLEITCDSTAMATNLMORQLIQTVAEKVGNIIETHEIRIFGPQAPS
MRKGPLHVKGRGRPDITG"
5354. 7489
/gene="Cg10006"
5354. 7489
/gene="Cg10006"
/EC number="5.99.1.3"
/note="PF00204:DNA topoisomerase II (N-terminal region)
PF00986:DNA gyrase B subunit, carboxyl terminus
PF01751:Toprim domain
PF02518:Histidine Kinase-, DNA gyrase B-, phycochrome-like
ATPase
TIGR01055:pare Greg: DNA topoisomerase IV, B subunit
TIGR01058:pare Gpos: DNA topoisomerase IV, B subunit
TIGR01059:gyrB: DNA gyrase, B subunit"
/codon_start=1
/transl_table=1
/product="DNA gyrase (topoisomerase II) B subunit"
/protein_id="BAB97399.1"
/db_xref="GI:21322770"
/translation="MRGTTGPKRVKRLRLRLISSECSLKVANTENHNDASSITILE
GLBAYRKPGMYIGSTGPRGLHLIMEVDNSVDENAGAHATKVEITLDDGQVVD
DGRGIPVDMHPSGAPVQVWMTOLHAGKPSDSYAGSLHGVGIVNALSRVVA
DIKHGKHWYONPEKSVDELREGNARGGTTTRFMPDAIEPTTEFDEPDSRLQ
EMALFNKGLITLTDNRATDELELBALBEGGTATLSLDEIDNETELVEETTDAR
KPKREKKRIHYENGLEDVHYLNRSKTNTHPSIVSEAGDHEVEVAMQNSYK
ESVHTFANTINTREGGTGEBGFSALTSLNKRVAREKTLKEKANATGDDQBSGA
VIVSVGDPQEGQTKTLGNTETKSFPORVANEHIHMEVLANEAKVILINAVSA
QARLAARARDLVRKRSATDGLGPKGLADCRSDPKSELYIVEGASGASGSD
SMFOALIEBRLKINVEKARLDKVLKNAVOALITAGTGIDHFDINKRYHIVLM
ADAPVDGHIATLTLLTFRMPDLNAGHYLAOPPLYLKNMGSGPRAVSEED
EQLNEGLAGARKINKDDGIGQYKGLGENMASELMTMDPIVRLIRVYDITDARADE
LFSIMGDDVVARNSFTIRNAKDVFFLDI"
7830. 8798
/gene="Cg10007"
7830. 8798
/gene="Cg10007"
/codon_start=1
/transl_table=1
/product="Hypothetical protein"
/protein_id="BAB97400.1"
/db_xref="GI:21322771"
/translation="MIOYILIKISLKNKAREMKTLITDITLGHVSTEBPTTAHY
LPDGGTPOVIOASDNKDSOLYMLMPGGMGYRPRPALANKAFVIAIGELRGG
OSSAKRSKSWGCHDASVDPLOIAAKALDIEBHPRPFSHSGGIGISCLFPA
RPEAKYNURALFVIGASGPPRPSPKMGKRLGIGVLTGIGIGHVGFPGKVLK
DIVYGRSGGTHREKRRFRKNSLDDITADINVEYMKKVSIPITPSRCDDDCP
OASIDGLAFVPAQIKMIEIPEALGHNRAKEPSTYKLEORL"
complement (8795. 9466)
/gene="Cg10008"
complement (8795. 9466)
/gene="Cg10008"
/codon_start=1
/transl_table=1
/product="Hypothetical membrane protein"
/protein_id="BAB97401.1"
/db_xref="GI:21322772"
/translation="MSNPVKLPDPAOFSLPTRAMFFVGIFIMSFALITVAGLG
TTDISLAPVMTAASGLSGWTIYNGMILCOIIVRSQKPMLOIIMAFJFG
ICDLSICLTMAOTDNFPAWITWIVSTLMSIGFIVLPITFIAGEIVSALVK
FNVFEGTKQIVDMTFVSAALISMTWIGLISVBEGIVRAAFIIGFPAVQMKRLYL
RSIGH"
complement (9471. 9914)
/gene="Cg10009"
complement (9471. 9914)
/gene="Cg10009"
/codon_start=1
/transl_table=1
/product="Hypothetical protein"
/protein_id="BAB97402.1"
/db_xref="GI:21322773"
/translation="MTSLRDLLVSTNADAAIADMSAFIEETVYKSGISGIALKGMG

AAKIDSDIVAKSRRLIPEIADSLDGLMOEYONGTRADPGALHANSSTALDAIIS
VADKNAESINVGKYYKGVGRGAARTIEGLPAISQLEKNAQ"
complement (10104. 11177)
/gene="Cg10010"
complement (10104. 11177)
/gene="Cg10010"
/codon_start=1
/transl_table=1
/product="Hypothetical protein"
/protein_id="BAB97403.1"
/db_xref="GI:21322774"
/translation="WFMAGRTLPFHIAPLPVARVPHQIRATDADRTSLQDLELS
Query Match 100.0%; Score 1271; DB 1; Length 340000;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1271; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGCGCATCAATGAATCTAGTTCATATATTTGACAAATAGCCTAGTGAAGTGCAGAAC 60
Db 276729 GGCGCATCAATGAATCTAGTTCATATATTTGACAAATAGCCTAGTGAAGTGCAGAAC 276788
QY 61 TGGCAACAAACTACCCGCAATTGTGTATGATTTAGTGTGCAAAAACGCAAGAT 120
Db 276789 TGGCAACAAACTACCCGCAATTGTGTATGATTTAGTGTGCAAAAACGCAAGAT 276848
QY 121 TCATTCAAGCTGGAGGTGTGGCATCCAGAGGAGCCCTGGAACGAGATGATTAAGTTA 180
Db 276849 TCATTCAAGCTGGAGGTGTGGCATCCAGAGGAGCCCTGGAACGAGATGATTAAGTTA 276908
QY 181 TGGCGCTACGAAATGCGCAAGTCTAAAAACCTCCCTGCTGCAAGTTTGGGCAATGA 240
Db 276909 TGGCGCTACGAAATGCGCAAGTCTAAAAACCTCCCTGCTGCAAGTTTGGGCAATGA 276968
QY 241 CCGGATGGATTTGGGCTGTGCTCTGTGTATTCATACGAGCTACGATGGTGGGACAC 300
Db 276969 CCGGATGGATTTGGGCTGTGCTCTGTGTATTCATACGAGCTACGATGGTGGGACAC 277028
QY 301 CCGCATGTTTCCGCGCTGATTTTGGCGGCTCCACCGAATGCTGTGATGCGCTCGT 360
Db 277029 CCGCATGTTTCCGCGCTGATTTTGGCGGCTCCACCGAATGCTGTGATGCGCTCGT 277088
QY 361 TGGGCGGAGGCGCCCTGGGCGCATTCGGCTACCAATGCTGTGATGCTGTGCGCA 420
Db 277089 TGGGCGGAGGCGCCCTGGGCGCATTCGGCTACCAATGCTGTGATGCTGTGCGCA 277148
QY 421 CGATTCATAGCGTTTCAATCCCGCTGATGTCAAAAAACCCCATTCGCCGTTCTA 480
Db 277149 CGATTCATAGCGTTTCAATCCCGCTGATGTCAAAAAACCCCATTCGCCGTTCTA 277208
QY 481 TTGGGTTTTGGCGCTTATGACGCAAGCTTACGAGTCACTGGGGCCAGGCGGAGGCTG 540
Db 277209 TTGGGTTTTGGCGCTTATGACGCAAGCTTACGAGTCACTGGGGCCAGGCGGAGGCTG 277268
QY 541 GTGCGGCTGGCACTTATCTCAATGCAAAATAGCGTTTCACTCTACTGTGATTTGGCGG 600
Db 277269 GTGCGGCTGGCACTTATCTCAATGCAAAATAGCGTTTCACTCTACTGTGATTTGGCGG 277328
QY 601 TTTCACCGAGTGGCGATGCGAGAGTTGATCTTTTGAATTAAGGCTCGAGTTCCG 660
Db 277329 TTTCACCGAGTGGCGATGCGAGAGTTGATCTTTTGAATTAAGGCTCGAGTTCCG 277388
QY 661 CTTTGGCTCTCTTTGTACAGCTGACGCTTGAATTCCTGCGCAACGAAAGAGATGCC 720
Db 277389 CTTTGGCTCTCTTTGTACAGCTGACGCTTGAATTCCTGCGCAACGAAAGAGATGCC 277448
QY 721 TTCTCTGCTCTGCGAGGTTTGAAGCTTACCAATGCTCTTGTGTAATTCAGGTCAGGC 780
Db 277449 TTCTCTGCTCTGCGAGGTTTGAAGCTTACCAATGCTCTTGTGTAATTCAGGTCAGGC 277508
QY 781 CCTATTGCGGCGCTGTGATCTTCTTGGGCTGTGACCATCGGTAATCTTCTTGGG 840
Db 277509 CCTATTGCGGCGCTGTGATCTTCTTGGGCTGTGACCATCGGTAATCTTCTTGGG 277568

QY 841 AAAGCTGCTAAATGACACTGATTTCTCTCTGATATCTCTCTGTTGTGCGATATGCGA 900
 Db 277569 AAAGCTGCTAAATGACACTGATTTCTCTCTGATATCTCTCTGTTGTGCGATATGCGA 277628
 QY 901 GTCAATTAATTTGGCGCTCCGGGCGGTTCCGTTCTTAATCCTTAAGCCCTCAAGTAATCA 960
 Db 277629 GTCAATTAATTTGGCGCTCCGGGCGGTTCCGTTCTTAATCCTTAAGCCCTCAAGTAATCA 277688
 QY 961 CAATTTGTGGGCAAAATGCGCATGTGATGTCGACAGGAATCTCTCCATTTGACCGCA 1020
 Db 277689 CAATTTGTGGGCAAAATGCGCATGTGATGTCGACAGGAATCTCTCCATTTGACCGCA 277748
 QY 1021 TCAAGCTTGGCAGAAATGCGCATGTGATGTCGACAGGAATCTCTCCATTTGACCGCA 277748
 Db 277749 TCAAGCTTGGCAGAAATGCGCATGTGATGTCGACAGGAATCTCTCCATTTGACCGCA 277808
 QY 1081 GCGATTACAGTGTGCGCATCTTCTTGCGCGTGCAGCGACCTTGTGAGCGTTGCGCT 1140
 Db 277809 GCGATTACAGTGTGCGCATCTTCTTGCGCGTGCAGCGACCTTGTGAGCGTTGCGCT 277868
 QY 1141 GCGACCACTGTTTGTGAGTGTGATCTTTCTTAATCTGATTAATTAACAAAT 1200
 Db 277869 GCGACCACTGTTTGTGAGTGTGATCTTTCTTAATCTGATTAATTAACAAAT 277928
 QY 1201 CCGATGCGCTCAATTTGAAGGGGATGCGGATTTTGAAGGACCTAGAAAAGCTTAAG 1260
 Db 277929 CCGATGCGCTCAATTTGAAGGGGATGCGGATTTTGAAGGACCTAGAAAAGCTTAAG 277988
 QY 1261 CAGACAGCGCT 1271
 Db 277989 CAGACAGCGCT 277999

RESULT 8
 AX120085
 LOCUS AX120085 349980 bp DNA linear PAT 11-MAY-2001
 DEFINITION Sequence 1 from Patent EP1108790.
 ACCESSION AX120085 AX114121
 VERSION AX120085.1 GI:14036800
 KEYWORDS
 SOURCE
 ORGANISM
 Corynebacterium glutamicum
 Corynebacterium glutamicum
 Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 Corynebacteriaceae; Corynebacteriaceae; Corynebacterium.
 REFERENCE
 AUTHORS
 TITLE Nakagawa, S., Mizoguchi, H., Ando, S., Hayashi, M., Ochiai, K.,
 Yokoi, H., Tateishi, N., Senoh, A., Ikeda, M. and Ozaki, A.
 JOURNAL Novel polynucleotides
 Patent: EP 1108790-A 1 20-JUN-2001;
 KYOMA HAKKO KOGYO CO., LTD. (JP)
 FEATURES
 source
 1. 349980
 /organism="Corynebacterium glutamicum"
 /mol_type="genomic DNA"
 /db_xref="taxon:1718"
 /note="Seq 1 to long (3,309,400) split in 11, seq 1
 0.000.001 0.349.980 349.980"

ORIGIN
 Query Match 100.0%; Score 1271; DB 6; Length 349980;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1271; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCGCGATCAATGATCTGCTTCAATATTTGCAATATGCTTAAGTGTGCGCAAC 60
 Db 276729 GCGCGATCAATGATCTGCTTCAATATTTGCAATATGCTTAAGTGTGCGCAAC 276788
 QY 61 TGGCAACAACGACCCGCGCAATGTGTGATGTTAGTGTGCAAAAAACGCAAGAT 120
 Db 276789 TGGCAACAACGACCCGCGCAATGTGTGATGTTAGTGTGCAAAAAACGCAAGAT 276848
 QY 121 TCATTACCGCTGAGGTGTGCGCATTCACGACCGCTTGAAACGATGATTAAGTTA 180

Db 276849 TCATTACCGCTGAGGTGTGCGCATTCACGACCGCTTGAAACGATGATTAAGTTA 276908
 QY 181 TGGCGCTACGAATGCGCAAGGTCTAATAAACCCTCCCTGCTGCGAGTTTGGCATGTA 240
 Db 276909 TGGCGCTACGAATGCGCAAGGTCTAATAAACCCTCCCTGCTGCGAGTTTGGCATGTA 276968
 QY 241 CCGATTTGATTTGGGTTGTGCTCTTGTGTTAATTAATAGGCTACGAATGTTGGGCGAC 300
 Db 276969 CCGATTTGATTTGGGTTGTGCTCTTGTGTTAATTAATAGGCTACGAATGTTGGGCGAC 277028
 QY 301 CCGATTTGATTTGGGTTGTGCTCTTGTGTTAATTAATAGGCTACGAATGTTGGGCGAC 360
 Db 277029 CCGATTTGATTTGGGTTGTGCTCTTGTGTTAATTAATAGGCTACGAATGTTGGGCGAC 277088
 QY 361 TGGGCGCAGCGCCCTTGGGCGCATGCGGCTACCAATGCTGTGTAACCTCGCGCA 420
 Db 277089 TGGGCGCAGCGCCCTTGGGCGCATGCGGCTACCAATGCTGTGTAACCTCGCGCA 277148
 QY 421 CGATTCTATGCGTTTCAATCCCGCTGATGTGTCATAAACCCTATGCGCGTTCTA 480
 Db 277149 CGATTCTATGCGTTTCAATCCCGCTGATGTGTCATAAACCCTATGCGCGTTCTA 277208
 QY 481 TTGCGTTTGGGCTTATGACGAAAGCTTACGATCTGCGGCGCAAGGCTCGAGCTG 540
 Db 277209 TTGCGTTTGGGCTTATGACGAAAGCTTACGATCTGCGGCGCAAGGCTCGAGCTG 277268
 QY 541 GTGCGGCGGCGACTTATCTCAATGCAATAGAGTTTCACTCTTCTGATTTGCGCGG 600
 Db 277269 GTGCGGCGGCGACTTATCTCAATGCAATAGAGTTTCACTCTTCTGATTTGCGCGG 277328
 QY 601 TCTACCGAGTGGCGCATGCGAGATGTAATCTTTGAATTAAGGCTCGAGTTGCG 660
 Db 277329 TCTACCGAGTGGCGCATGCGAGATGTAATCTTTGAATTAAGGCTCGAGTTGCG 277388
 QY 661 CTTTGTCTCTCTTTGTGACGCTGACTTGTGATCTCTCCCGAAGCAAAAGCATGCC 720
 Db 277389 CTTTGTCTCTCTTTGTGACGCTGACTTGTGATCTCTCCCGAAGCAAAAGCATGCC 277448
 QY 721 TTCTGTGCTGCTGCGAGTTTGAAGCTTCAACATGCTCTGTGTGTAATTCAGAGTGC 780
 Db 277449 TTCTGTGCTGCTGCGAGTTTGAAGCTTCAACATGCTCTGTGTGTAATTCAGAGTGC 277508
 QY 781 CCAATTTGGGCGCTGCTATCTTCTTGGGCTGTGACCATCCGCTACTCTTCTTGGG 840
 Db 277509 CCAATTTGGGCGCTGCTATCTTCTTGGGCTGTGACCATCCGCTACTCTTCTTGGG 277568
 QY 841 AAAGCTGCTAATGACAACTGATTTCTCCTGATTTCTCTTGTGCGAGTATGCA 900
 Db 277569 AAAGCTGCTAATGACAACTGATTTCTCCTGATTTCTCTTGTGCGAGTATGCA 277628
 QY 901 GTCAATTAATTTGGCGCTCCGGGCGGTTCCGTTCTTAATCCTTAAGCCCTCAAGTAATCA 960
 Db 277629 GTCAATTAATTTGGCGCTCCGGGCGGTTCCGTTCTTAATCCTTAAGCCCTCAAGTAATCA 277688
 QY 961 CAATTTGTGGGCAAAATGCGCATGTGATGTCGACAGGAATCTCTCCATTTGACCGCA 1020
 Db 277689 CAATTTGTGGGCAAAATGCGCATGTGATGTCGACAGGAATCTCTCCATTTGACCGCA 277748
 QY 1021 TCAAGCTTGGCAGAAATGCGCATGTGATGTCGACAGGAATCTCTCCATTTGACCGCA 277748
 Db 277749 TCAAGCTTGGCAGAAATGCGCATGTGATGTCGACAGGAATCTCTCCATTTGACCGCA 277808
 QY 1081 GCGATTACAGTGTGCGCATCTTCTTGCGCGTGCAGCGACCTTGTGAGCGTTGCGCT 1140
 Db 277809 GCGATTACAGTGTGCGCATCTTCTTGCGCGTGCAGCGACCTTGTGAGCGTTGCGCT 277868
 QY 1141 GCGACCACTGTTTGTGAGTGTGATCTTTCTTAATCTGATTAATTAACAAAT 1200
 Db 277869 GCGACCACTGTTTGTGAGTGTGATCTTTCTTAATCTGATTAATTAACAAAT 277928
 QY 1201 CCGATGCGCTCAATTTGAAGGGGATGCGGATTTTGAAGGACCTAGAAAAGCTTAAG 1260
 Db 277929 CCGATGCGCTCAATTTGAAGGGGATGCGGATTTTGAAGGACCTAGAAAAGCTTAAG 277988

QY 1261 CAGACAGCGCT 1271
 Db 277989 CAGACAGCGCT 277999

RESULT 9
 AF454053/c 2105 bp DNA linear BCT 27-JUN-2002

LOCUS AF454053/c
 DEFINITION Corynebacterium glutamicum BrnF (brnF), BrnF (brnF), and Lrp-like
 regulator (Lrp) gene, complete cds.

ACCESSION AF454053
 VERSION AF454053.1 GI:21311379

KEYWORDS
 SOURCE

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 MEDLINE
 PUBMED
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL

Export of L-isoleucine from Corynebacterium glutamicum: a
 two-gene-encoded member of a new translocator family
 J. Bacteriol. 184 (14), 3947-3956 (2002)
 22077265
 12081967
 2 (bases 1 to 2105)
 Kennerknecht, N., Eggeling, L., and Sahm, H.
 Direct Submission
 Submitted (07-NOV-2001) IBT-1, Forschungszentrum, Leo-Brandt Str.,
 Juelich 52425, Germany

FEATURES
 source

Location/Qualifiers
 1..2105
 /organism="Corynebacterium glutamicum"
 /mol_type="genomic DNA"
 /db_xref="taxon:1718"
 complement(385..711)
 /gene="brnF"
 complement(385..711)
 /note="Lrp-E transporter component 1; branched-chain amino
 acid export protein"
 /codon_start=1
 /transl_table=11
 /product="BrnF"
 /protein_id="AA046685.1"
 /db_xref="GI:21311380"
 /translation="MTTDFSCILVAVCAVTFALRAVPLILKPLESGQFVGKAM
 NMPAGILALITASTFRSNAIDKTLTFGLIAVAIVVAHLGGRTLLSVAGITIVF
 GVNLP"
 complement(708..1463)
 /gene="brnF"
 complement(708..1463)
 /note="Lrp-E transporter component 2; branched-chain amino
 acid export protein"
 /codon_start=1
 /transl_table=11
 /product="BrnF"
 /protein_id="AA046686.1"
 /db_xref="GI:21311381"
 /translation="MGTDLHSSLEVSFSKALPEDDKGYRRIEAGLQKTSLAAL
 GMPPIAGLAVLYGYEMWAPLFGSLIFAGSTEMVIAVAGAAPGALATLTL
 VNRHVFVAFSPILHVNKNPILAFVSPALIDEAVAVTARAGSAMPILSMOAFH
 SYWFGALGVAAELIPEIKLEBALCSLFTVTLDSCKTKKQIPSLIAGLSFTI
 ALVIRGQALFALLIFGLITIRYFLKAK"
 1538..1993
 /gene="lrp"
 1538..1993
 /gene="lrp"
 /note="Lrp"
 /codon_start=1
 /transl_table=11

ORIGIN

Query Match 99.1%; Score 1260; DB 1; Length 2105;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1271; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

/product="Lrp-like regulator"
 /protein_id="AA046687.1"
 /db_xref="GI:21311382"
 /translation="MTLSDIDRALIIEISANRISLALADKYLTPGCBLEHRLRG
 ASGIIIGSADHPAVNRGFEYVTVLNSFDRSTVDNFESSVAGHDEVLHRLRG
 SPDYFVRIGVADLEAYEPLSSHIQTPGIATISRFKRVKVPAPQV"

QY 1 GCGGATCAATGGAATCTAGCTTCAATATTTGCAATAGCTAGTTAGAGTGGCAAC 60
 Db 1563 GCGGATCAATGGAATCTAGCTTCAATATTTGCAATAGCTAGTTAGAGTGGCAAC 1504
 QY 61 TGGCAACAAATCAACCGGCAATTTGTGTATGTTAGTGTGCAAAAACGCAAGAT 120
 Db 1503 TGGCAACAAATCAACCGGCAATTTGTGTATGTTAGTGTGCAAAAACGCAAGAT 1444
 QY 121 TCATTCAGGCTGAGGTGTGCGCATCCAGGAGCGCTGGAACCAATGATTAAGGTTA 180
 Db 1443 TCATTCAGGCTGAGGTGTGCGCATCCAGGAGCGCTGGAACCAATGATTAAGGTTA 1384
 QY 181 TCGGCGCTAGGAATCGCGCAAGGTCTAAAAACCTGCTGCGAGGTTGGGCAATGA 240
 Db 1383 TCGGCGCTAGGAATCGCGCAAGGTCTAAAAACCTGCTGCGAGGTTGGGCAATGA 1324
 QY 241 CCGGATTTGATTCGTTGTGTGTGTGTATTCATACGGCTAGCAATGTGTGGCAGC 300
 Db 1323 CCGGATTTGATTCGTTGTGTGTGTGTATTCATACGGCTAGCAATGTGTGGCAGC 1264
 QY 301 CCGCATGTTTCCGCGCTGATTTTCCGCGGCTCACCGCAATGCTGTCTCCCTCCG 360
 Db 1263 CCGCATGTTTCCGCGCTGATTTTCCGCGGCTCACCGCAATGCTGTCTCCCTCCG 1204
 QY 361 TGTGGGCGAGCGCCCTGCGGCGCATCGGCTCACCAATGCTGTGTGTAATTCGCGCA 420
 Db 1203 TGTGGGCGAGCGCCCTGCGGCGCATCGGCTCACCAATGCTGTGTGTAATTCGCGCA 1144
 QY 421 GGAATTCATGCGTTTCAATCCCGGTGATGTGTGTGCAAAAACCCATTCGCCGTTCTA 480
 Db 1143 GGAATTCATGCGTTTCAATCCCGGTGATGTGTGTGCAAAAACCCATTCGCCGTTCTA 1084
 QY 481 TTGGGTTTGGCGGCTTATCGACGAAGCTTACGAGTCACTGCGGCGCAGCGCGAGCTG 540
 Db 1083 TTGGGTTTGGCGGCTTATCGACGAAGCTTACGAGTCACTGCGGCGCAGCGCGAGCTG 1024
 QY 541 GTGCGGCTGGGCACTTATCTCATATGCAATAGCGTTTCACTCTGAGTATTCGGCGG 600
 Db 1023 GTGCGGCTGGGCACTTATCTCATATGCAATAGCGTTTCACTCTGAGTATTCGGCGG 964
 QY 601 TCTACCGGAGTGGCGATTCGACAGTGTATTCCTTTGAATTAAGGCGCTCGAGTTGCG 660
 Db 963 TCTACCGGAGTGGCGATTCGACAGTGTATTCCTTTGAATTAAGGCGCTCGAGTTGCG 904
 QY 661 CTTTGGCTCTCTCTTTGCAAGCTGACTTGGATTCCTGCGGACGAAAAAGCAATCCC 720
 Db 903 CTTTGGCTCTCTCTTTGCAAGCTGACTTGGATTCCTGCGGACGAAAAAGCAATCCC 844
 QY 721 TTTCTGTGTGTGCGAGTTTGAAGCTTCAATGCTCTGTGTGTATTCAGGTCAGGC 780
 Db 843 TTTCTGTGTGTGCGAGTTTGAAGCTTCAATGCTCTGTGTGTATTCAGGTCAGGC 764
 QY 781 CTTATTTGCGGCGCTGCTGATCTTTCTTGGGCTGTGCAATCGGATCTTCTTGGG 840
 Db 763 CTTATTTGCGGCGCTGCTGATCTTTCTTGGGCTGTGCAATCGGATCTTCTTGGG 724
 QY 841 AAAGGCTGCTAAATGACATGATTTTCTCTGATTTCTCTTGTGTGCGAGTATGCA 900
 Db 723 AAAGGCTGCTAAATGACATGATTTTCTCTGATTTCTCTTGTGTGCGAGTATGCA 664
 QY 901 GTCAATTAATTTGCGGCTCGGCGGTTCCGTTCTTAATTCCTTAAGCCCTAGTGAATCA 960

```

Db      663  GTCAATTAATTGGCGCTCCGGGCGGCTCGTTCTTAATCTTAAAGCCCTACGTAATCA 604
Qy      961  CAATTTGGGGCAAAATGGCGATGGGATGGCAGAGGAATCTTGCAATTTTACCGCA 1020
Db      603  CAATTTGGGGCAAAATGGCGATGGGATGGCAGAGGAATCTTGCAATTTTACCGCA 544
Qy      1021  TCAACGTTTCGAGCAATGCGATGATCTGAAGACTTACCTTGGTCTGATTCGCGGT 1080
Db      543  TCAAGTTTTCGAGCAATGCGATGATCTGAAGACTTACCTTGGTCTGATTCGCGGT 484
Qy      1081  GCGATTACAGTGTGGCGCATTTTCTTGCGGCTGCAAGCAGCTTGTGAGGCTTGCGCT 1140
Db      483  GCGATTACAGTGTGGCGCATTTTCTTGCGGCTGCAAGCAGCTTGTGAGGCTTGCGCT 424
Qy      1141  GGCACCATCGTTTGTGGAAGTGTGAATCTTTCTTAAACTGATTAATACAAAAT 1200
Db      423  GGCACCATCGTTTGTGGAAGTGTGAATCTTTCTTAAACTGATTAATACAAAAT 364
Qy      1201  CCGCATGCCCTCAATTTGAAGGGGATGCGGATTTTGAAGAACCTAG-AAAAGCTTAA 1259
Db      363  CCGCATGCCCTCAATTTGAAGGGGATGCGGATTTTGAAGAACCTAGAAAAGGCTTAA 304
Qy      1260  GCGAGCAGGCGCT 1271
Db      303  GCGAGCAGGCGCT 292

RESULT 10
AR31954 753 bp DNA linear PAT 18-DEC-2003
LOCUS AR31954
DEFINITION Sequence 2 from patent US 6613545.
ACCESSION AR31954
VERSION AR31954.1 GI:40115725
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 753)
AUTHORS Kemerhnecht,N., Salm,H., Eggeling,L. and Pfeifferle,W.
TITLE Nucleotide sequences coding for the export of branched chain amino
JOURNAL Patent: US 6613545-A 2 02-SBP-2003;
FEATURES
SOURCE location/Qualifiers
1..753
/organism="Unknown"
/mol_type="genomic DNA"

ORIGIN
Query Match 59.2%; Score 753; DB 6; Length 753;
Best Local Similarity 100.0%; Pred. No. 3.3e-196;
Matches 753; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 101 GTGCAAAAACGCAAGATTCATTCAAGCTGAGGTGCGCAATCCAGGACACCTG 160
Db 1 GTGCAAAAACGCAAGATTCATTCAAGCTGAGGTGCGCAATCCAGGACACCTG 60
Qy 161 GAACCATGATTAAGGTATCGGCGCTACGAAATGCGCAAGTCTAAAAACCTCCCT 220
Db 61 GAACCATGATTAAGGTATCGGCGCTACGAAATGCGCAAGTCTAAAAACCTCCCT 120
Qy 221 GCTGCAAGTTTGGGCGATGATCCGATGATGATGCGTTGTCTCTGTTTCAATAC 280
Db 121 GCTGCAAGTTTGGGCGATGATCCGATGATGATGCGTTGTCTCTGTTTCAATAC 180
Qy 281 GGTACAGATGATGAGGCGAGCCCACTGTTTCCGGCTGATTTTGGCGGGCTCCACGGA 340
Db 181 GGTACAGATGATGAGGCGAGCCCACTGTTTCCGGCTGATTTTGGCGGGCTCCACGGA 240
Qy 341 ATGCTGATCGCCCTCGTTGTGGGCGAGCGCCCTGCGGCGCATCGCGCTCACACA 400
Db 241 ATGCTGATCGCCCTCGTTGTGGGCGAGCGCCCTGCGGCGCATCGCGCTCACACA 300

```

```

Qy 401 TTGCTGTAAGTCTCCGACGATTCATGCGTTTTCATTCGCCCTGCATGCTGCAAA 460
Db 301 TTGCTGTAAGTCTCCGACGATTCATGCGTTTTCATTCGCCCTGCATGCTGCAAA 360
Qy 461 AACCCATGGCCGTTTCTATTCGGTTTTCGGCTTATGAGAAAGCCATACGAGTCACT 520
Db 361 AACCCATGGCCGTTTCTATTCGGTTTTCGGCTTATGAGAAAGCCATACGAGTCACT 420
Qy 521 GCGGCGAGCCGCGAGGCTGCTGCGGCGTGCAGCACTTATCTCAATGCAAAATAGGTTTCA 580
Db 421 GCGGCGAGCCGCGAGGCTGCTGCGGCGTGCAGCACTTATCTCAATGCAAAATAGGTTTCA 480
Qy 581 TCTTACTGAGTATTCGCGCGCTCTCACCGAGTGGCATGCGAGATGATCTTTTGA 640
Db 481 TCTTACTGAGTATTCGCGCGCTCTCACCGAGTGGCATGCGAGATGATCTTTTGA 540
Qy 641 ATTAAAGGCTCGAGTTCGCGCTTCTCTCTTGTGACGCTGATTCGATTCCTG 700
Db 541 ATTAAAGGCTCGAGTTCGCGCTTCTCTCTTGTGACGCTGATTCGATTCCTG 600
Qy 701 CGAACGAAAAGCAGATCCCTTCTCTGCTGCTGCAAGTTTGAAGTTTCACTTCTT 760
Db 601 CGAACGAAAAGCAGATCCCTTCTCTGCTGCTGCAAGTTTGAAGTTTCACTTCTT 660
Qy 761 GTGGTAATTCAGATCAGGCGCTTATTTGCGGCGCTGCTGATTCCTTGGGCTGTGACC 820
Db 661 GTGGTAATTCAGATCAGGCGCTTATTTGCGGCGCTGCTGATTCCTTGGGCTGTGACC 720
Qy 821 ATCCGTAATTCCTTCTTGGGAAAGGCTCTAA 853
Db 721 ATCCGTAATTCCTTCTTGGGAAAGGCTCTAA 753

RESULT 11
AX120373 753 bp DNA linear PAT 11-MAY-2001
LOCUS AX120373
DEFINITION Sequence 289 from Patent EP1108790.
ACCESSION AX120373
VERSION AX120373.1 GI:14037088
KEYWORDS
SOURCE Corynebacterium glutamicum
ORGANISM Corynebacterium glutamicum
REFERENCE 1
AUTHORS Nakagawa,S., Mizoguchi,H., Ando,S., Hayashi,M., Ochiai,K.,
TITLE Novel polynucleotides
JOURNAL Patent: EP 1108790-A 289 20-JUN-2001;
JOURNAL KOWA HAKKO KOGYO CO., LTD. (JP)
FEATURES
SOURCE location/Qualifiers
1..753
/organism="Corynebacterium glutamicum"
/mol_type="unassigned DNA"
/db_xref="taxon:1718"

ORIGIN
Query Match 59.2%; Score 753; DB 6; Length 753;
Best Local Similarity 100.0%; Pred. No. 3.3e-196;
Matches 753; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 101 GTGCAAAAACGCAAGATTCATTCAAGCTGAGGTGCGCAATCCAGGACACCTG 160
Db 1 GTGCAAAAACGCAAGATTCATTCAAGCTGAGGTGCGCAATCCAGGACACCTG 60
Qy 161 GAACCATGATTAAGGTATTCGCGCTACGAAATGCGCAAGTCTAAAAACCTCCCT 220
Db 61 GAACCATGATTAAGGTATTCGCGCTACGAAATGCGCAAGTCTAAAAACCTCCCT 120
Qy 221 GCTGCAAGTTTGGGCGATGATCCGATGATGATGCGTTGTCTCTGTTTCAATAC 280
Db 121 GCTGCAAGTTTGGGCGATGATCCGATGATGATGCGTTGTCTCTGTTTCAATAC 180

```



```

REFERENCE 1 (bases 1 to 753)
AUTHORS Kemerukunento,N., Salm,H., Eggering,L. and Pfeifferle,M.
TITLE Nucleotide sequence encoding the discharge of branched amino acid,
JOURNAL method of isolating the same and utilization thereof
DEUSSA HUELS AG,FORSCHUNGSZENTRUM JUELICH GMBH
OS Corynebacterium glutamicum ATCC14752
PN JP 2001169788-A/2
PD 26-JUN-2001
PF 24-OCT-2000 JP 2000324315
PR 27-OCT-1999 DE 19951708.8
PI NICOLE KEMERUKUNENTO,HERMANN SALM,LOTHAR EGGERING,WALTER PI
PFEIFFERLE
PC C12N15/09,C07K14/34,C12N1/21,C12P13/06,C12P13/06,C12P13/08//
PC C12N15/09,C12R1:15),(C12N1/21,C12R1:15),(C12P13/06,C12R1:15),
PC C12N15/00,C12R1:15)
PC C12N15/00,C12R1:15)
CC bnf
CC key
FH key
FT CDS Location/Qualifiers
1..753
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"

FEATURES
source
1..753
Location/Qualifiers
1..753
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"

ORIGIN
Query Match 59.2%; Score 753; DB 6; Length 753;
Best Local Similarity 100.0%; Pred. No. 3,3e-196; Mismatches 0; Indels 0; Gaps 0;
Matches 753; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

101 GTGCAAAAACGCAAGATTCATTCAAGCTGAGGTGTGCCATCCAGGAGCCCTG 160
DB 1 GTGCAAAAACGCAAGATTCATTCAAGCTGAGGTGTGCCATCCAGGAGCCCTG 60
161 GAACCAAGATGATTAAGTTATCGGGCTACGAATCCGGCAAGGCTTAAACCTCCCT 220
DB 61 GAACCAAGATGATTAAGTTATCGGGCTACGAATCCGGCAAGGCTTAAACCTCCCT 120
221 GCTGCAAGTTGGGATGATACCCGATGATTTGGTGTCTCTTGTATTAATCAATAC 280
DB 121 GCTGCAAGTTGGGATGATACCCGATGATTTGGTGTCTCTTGTATTAATCAATAC 180
281 GGTTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 340
DB 181 GGTTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 240
341 ATGCTGATCATGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 400
DB 241 ATGCTGATCATGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 300
401 TTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 460
DB 301 TTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 360
461 AACCCATTGACCGCTTCTGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTC 520
DB 361 AACCCATTGACCGCTTCTGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTC 420
521 GCGGCAAGGCGGCAAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 580
DB 421 GCGGCAAGGCGGCAAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
581 TCCCTAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 640
DB 481 TCCCTAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 540
641 ATTAAGGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 700
DB 541 ATTAAGGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
701 GGAAGGAAAGCAATCCCTTCTGCTGCTGCGAGGTTGAAGTTGAACATTCCTCTT 760

```

```

DB 601 CGACGAAAAGCAATCCCTTCTGCTGCTGCGAGGTTGAAGCTTCAACATTCCTT 660
QY 761 GTGTAAATTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 820
DB 661 GTGTAAATTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
QY 821 ATCCGATCTCTCTCTGAGAAAGGCTGTAA 853
DB 721 ATCCGATCTCTCTCTGAGAAAGGCTGTAA 753

RESULT 14
BD162490
LOCUS Novel polynucleotide. 753 bp DNA linear PAT 17-JAN-2003
DEFINITION BD162490
ACCESSION BD162490.1 GI:27868248
VERSION JP 2002191370-A/289.
KEYWORDS unclassified
SOURCE unclassified
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 753)
AUTHORS Nakagawa,S., Mizoguchi,H., Ando,S., Hayashi,M., Ochiai,K.,
TITLE Yokoi,H., Tateishi,N., Senoo,A., Ikeda,K. and Ozaki,A.
JOURNAL Novel polynucleotide
PATENT: JP 2002191370-A 289 09-JUL-2002;
KYOMA HAKKO KOGYO CO LTD
OS Corynebacterium glutamicum
PN JP 2002191370-A/289
PD 09-JUL-2002 JP 2000405096
PF 15-DEC-2000 JP 2000405096
PI SATOSHI NAKAGAWA,HIROSHI MIZOGUCHI,SEIKO ANDO,MIKIO HAYASHI,
PI KEIKO OCHIAI,
PI HAHUHIKO YOKOI,NAOKO TATEISHI,AKIHIRO SENOO,MASATO IKEDA,AKIO
PI OZAKI
PC C12N15/09,C12N15/09,C07K14/34,C07K16/12,C07K16/40,C12M1/00,PC
PC C12N1/15,
PC C12N1/19,C12N1/21,C12N5/10,C12N9/00,C12N9/02,C12P7/40,C12P13/
PC 04,C12P13/08,
PC C12P19/00,C12P19/34,C12P21/02,C12Q1/37,C12Q1/68,G01N33/53,PC
G01N33/56,
PC G01N33/56,G01N33/58,G01N37/00//C12P21/08,(C12N1/21,C12R1:15),
PC (C12N1/21,C12R1:13),(C12N1/21,C12R1:01),(C12P13/08,C12R1:15),
PC C12N15/00,
PC C12N5/00,C12N15/00
CC Novel polynucleotide
FH key
FT source
1..753
/organism="Corynebacterium glutamicum".

FEATURES
source
1..753
Location/Qualifiers
1..753
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"

ORIGIN
Query Match 59.2%; Score 753; DB 6; Length 753;
Best Local Similarity 100.0%; Pred. No. 3,3e-196; Mismatches 0; Indels 0; Gaps 0;
Matches 753; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

101 GTGCAAAAACGCAAGATTCATTCAAGCTGAGGTGTGCCATCCAGGAGCCCTG 160
DB 1 GTGCAAAAACGCAAGATTCATTCAAGCTGAGGTGTGCCATCCAGGAGCCCTG 60
161 GAACCAAGATGATTAAGTTATCGGGCTACGAATCCGGCAAGGCTTAAACCTCCCT 220
DB 61 GAACCAAGATGATTAAGTTATCGGGCTACGAATCCGGCAAGGCTTAAACCTCCCT 120
221 GCTGCAAGTTGGGATGATACCCGATGATTTGGTGTCTCTTGTATTAATCAATAC 280
DB 121 GCTGCAAGTTGGGATGATACCCGATGATTTGGTGTCTCTTGTATTAATCAATAC 180
281 GGTTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 340

```

Db	181	GGCTACGAATGATGGGACAGCCCACTGTTTTCCGAGCTGATTTTCGGGGGCTCCACGAA	240
Qy	341	ATGCTGATCATGCCCCCTCGTTGTGGGCGAGAGGCCCCCTGGGGCGCATCGGGCTCACCA	400
Db	241	ATGCTGATCATGCCCCCTCGTTGTGGGCGAGAGGCCCCCTGGGGCGCATCGGGCTCACCA	300
Qy	401	TTGCTGGTGAATTCGCGCACGATATCTATGCGTTTCATTCCCGCTGCATGTGTCAAA	460
Db	301	TTGCTGGTGAATTCGCGCACGATATCTATGCGTTTCATTCCCGCTGCATGTGTCAAA	360
Qy	461	AACCCCATGCCCCGTTTCTATTCGTTCGCGTTTACGAGCCTTACGAGCTCACT	520
Db	361	AACCCCATGCCCCGTTTCTATTCGTTCGCGTTTACGAGAGCCTTACGAGCTCACT	420
Qy	521	GGCGGCGAGGCGCGCAGGCTGATGAGCGATGCGGACCTTACTAATGCAATATGAGGTTTCA	580
Db	421	GGCGGCGAGGCGCGCAGGCTGATGAGCGATGCGGACCTTACTAATGCAATATGAGGTTTCA	480
Qy	581	TCTTACTGAGTATTCGCGCGTCTCAACGAGATGCGCATCGAGAGTTGATTCCTTTGAA	640
Db	481	TCTTACTGAGTATTCGCGCGGTCTCAACGAGATGCGCATCGAGAGTTGATTCCTTTGAA	540
Qy	641	ATTAAAGGCGCTCGAGTTTCGCGCTTGTCTCTCTTTTGACAGCTGACTTTGATTCCTGC	700
Db	541	ATTAAAGGCGCTCGAGTTTCGCGCTTGTCTCTCTTTTGACAGCTGACTTTGATTCCTGC	600
Qy	701	CGAAGCAAAAAGAGATCCCTTCTGCGTCTCGAGGTTTGAGACTTCAACATTCGCTT	760
Db	601	CGAAGCAAAAAGAGATCCCTTCTGCGTCTCGAGGTTTGAGACTTCAACATTCGCTT	660
Qy	761	GTCGTAAATTCAGATCAGGCGCTTATTTCGCGCGCTGCTGATCTTCTTGAGTCTGTGACC	820
Db	661	GTCGTAAATTCAGATCAGGCGCTTATTTCGCGCGCTGCTGATCTTCTTGAGTCTGTGACC	720
Qy	821	ATCCGAGTACTCTTCTTGAGAAAGGCTGTAA	853
Db	721	ATCCGAGTACTCTTCTTGAGAAAGGCTGTAA	753

```

RESULT 15
AX137083/c
LOCUS
DEFINITION Sequence 7 from Patent EP1090993.
ACCESSION AX137083
VERSION AX137083.1 GI:14273428
KEYWORDS
SOURCE
ORGANISM
Corynebacterium glutamicum
Corynebacterium glutamicum
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacterineae; Corynebacteriaceae; corynebacterium.
REFERENCE
AUTHORS Moeckel,B., Pfefferle,W., Puenhler,A., Kalinowski,T. and Bathe,B.
TITLE Nucleotide sequences coding for the lrp gene
JOURNAL Patent: EP 1090993-A 7 11-APR-2001;
Degussa AG (DE)
FEATURES
source
1..397
Location/Qualifiers
/organism="Corynebacterium glutamicum"
/mol_type="unassigned DNA"
/db_xref="taxon:1718"
/note="lrp-Teil 1"
ORIGIN

```

Query Match	28.8%	Score 365.6;	DB 6;	Length 397;
Best Local Similarity	98.9%	Pred. No. 2e-89;		
Matches 368;	Conservative	0;	Mismatches 4;	Indels 0;
				Gaps 0

[illegible]

Db	316	AAACTACCCGCGCAATTGTGTGATGATTGTAGTGTGCAAAAAACCCAGAGATTCATTCMAA	25
Qy	129	GCGTCGAGGTTGTGCGCATCCAGAGCGCCCTGGAAACGATGATPAAAGTTATCGCGCT	188
Db	256	GCTCGAGAGTGTGCCATCCAGGCGACGCCCTGGAAACGATGATPAAAGTTATCGCGCT	197
Qy	189	ACGAATCGCGCAAGCTCTAAAAACCTCCCTGTGCGAGTTTGGGATATGACCCGATTG	248
Db	196	ACGAATCGCGCAAGGCTCTAAAACCTCCCTGTGCGAGTTTGGGCAATGACCCGATTG	137
Qy	249	GATATGCGTTTGTCTCTTGTATTCAATACGCGTACGAATGTTGGCGACGCCACTGT	308
Db	136	GATATGCGTTTGTCTCTTGTATTCAATACGCGTACGAATGTTGGCGACGCCACTGT	77
Qy	309	TTATCGGCGCTGATTTTCGCGGCGCTCACCGAATCTGTGATCGCCTCGTTTGGCG	368
Db	76	TTTCCGCGCTGATTTTCGCGGCGCTCACCGAATCTGTGATCGCCTCGTTTGGCG	17
Qy	369	CAGCGCCCCGTG	380
Db	16	CAGCGCCCCGTG	5

Search completed: April 19, 2004, 18:29:50
Job time : 5394 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 19, 2004, 10:56:53 ; Search time 591.9 Seconds
(without alignments)
9122.247 Million cell updates/sec

Title: US-10-608-504-1
Perfect score: 1271
Sequence: 1 gcgcgcatacgaatgacatcag.....aggcttaagcagacagcgcct 1271

Scoring table: IDENTITY NUC
Gapop 10.0, Gapept 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_25Jan04:*
1: geneseqn19808:*
2: geneseqn19908:*
3: geneseqn20008:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2003as:*
8: geneseqn2003bs:*
9: geneseqn2003cs:*
10: geneseqn2004as:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1271	100.0	1271	AAH21109	Aah21109 C. glutam
2	1271	100.0	1271	AAH21112	Aah21112 C. glutam
3	1271	100.0	349980	AAH64966	Aah64966 C. glutam
4	753	59.2	753	AAH21110	Aah21110 C. glutam
5	753	59.2	753	AAH65254	Aah65254 C. glutam
6	553	43.5	564	ACA01969	ACA01969 C. glutam
7	365.6	28.8	397	AAH61693	Aah61693 C. glutam
8	365.6	28.8	778	AAH61695	Aah61695 C. glutam
9	324	25.5	324	AAH21111	Aah21111 C. glutam
10	324	25.5	324	AAH65255	Aah65255 C. glutam
11	324	25.5	324	ACA01970	ACA01970 C. glutam
12	185	14.6	715	AAH61688	Aah61688 C. glutam
13	177	13.9	177	ACA01968	ACA01968 C. glutam
14	127.4	10.0	711	AAH68373	Aah68373 C. glutam
15	127.4	10.0	711	ACA01294	ACA01294 C. glutam
16	127.4	10.0	309400	AAH68534	Aah68534 C. glutam
17	48.2	3.8	2000	ADA71938	Ada71938 Rice gene
18	48	3.8	10732	AAH10594	Aa10594 Gene enco
19	40.8	3.2	7645	ADBS59610	Adbs59610 Human gen
20	40.2	3.2	4947	AAQ86927	AAQ86927 TSPS17 po
21	39.8	3.1	768	ADC91639	Adc91639 E. faeciu
22	39.2	3.1	6741	AAH10595	Aa10595 Gene enco
23	38.2	3.0	946	AAH55683	Aah55683 Human bre

24	37.6	3.0	759	7	ACF69125	Acf69125 Photorhab
25	37.6	3.0	5059	2	AAH84332	Aah84332 Stealth v
26	37.6	3.0	110000	7	ACF67367_20	Continuation (21 o
27	37.6	3.0	243072	7	ACF65382	ACF65382 Photorhab
28	37.2	2.9	1024	6	AAH66366	Aah66366 Helicobac
29	37.2	2.9	2000	7	ADA71938	Ada71938 Rice gene
30	36.8	2.9	1230	7	ACA42205	ACA42205 Prokaryot
31	36.4	2.9	751	2	AAV58489	Aav58489 3' fragme
32	36.4	2.9	751	2	AAV61146	Aav61146 3' CDNA s
33	36.4	2.9	751	3	AAA06252	AAA06252 Human imm
34	36.4	2.9	751	3	AB571157	Ab571157 Human pro
35	36.4	2.9	751	4	AAH93368	Aah93368 Human pro
36	36.4	2.9	751	4	AAH63460	Aah63460 Human pro
37	36.4	2.9	751	4	AAH02433	Aah02433 Prostate
38	36.4	2.9	751	4	AAH84682	Aah84682 Human pro
39	36.4	2.9	751	5	ACA59269	ACA59269 Prostate
40	36.4	2.9	751	5	AAH10011	Aah10011 Human pro
41	36.4	2.9	751	6	AB194832	Ab194832 Human NI-
42	36.4	2.9	751	6	AB585841	Ab585841 Prostate
43	36.4	2.9	751	7	ACC94996	Acc94996 Prostate
44	36.4	2.9	751	9	ADB13462	Adb13462 Human pro
45	36.2	2.8	19012	2	AAQ36952	Aaq36952 HSA gene.

ALIGNMENTS

RESULT 1
ID AAH21109 standard; DNA; 1271 BP.

AC AAH21109;
DT 05-SEP-2001 (first entry)
XX C. glutamicum DNA encoding brnF and brnE.

KW L-amino acid production; brnF; brnE; branched-chain amino acid;
KW Corynebacterium; leucine; isoleucine; valine; medicine;
KW animal nutrition; ds.

OS Corynebacterium glutamicum.

XX Key Location/Qualifiers
FH CDS 101..856
FT /tag= a "brnF"
FT /product= "brnF"
FT CDS 853..1179
FT /tag= b
FT /product= "brnE"

XX EP1096010-A1.

XX 02-MAY-2001.

XX 11-OCT-2000; 2000EP-0012057.

XX 27-OCT-1999; 99DE-01051708.

XX (DEGS) DEGUSSA AG.
XX (KERJ) FORSCHUNGSZENTRUM UEBELICH GMBH.

XX Kennernknecht N, Eggeling L, Sahn H, Pfeifferle W;
XX WPI; 2001-391595/42.

XX P-P-SDB; AAB66247, AAB66248.

XX New export genes from corynebacterium, useful for increasing
XX fermentative production of branched-chain amino acids.

XX Claim 4 (1); Page 13; 23pp; German.

XX This invention describes a novel isolated polynucleotide (I) containing

at least one sequence that (1) is 70% identical with a sequence that encodes at least one of 251 amino acid (aa) (3) or 108 aa (5) polypeptides; (1i) encodes a polypeptide at least 70% identical with (3) or (5); (1ii) is the complement of (1) or (1i), or (1iv) contains at least 15 consecutive bases from (1)-(1ii). The invention also describes (a) protein derived from sequences of 753 bp (2) or 324 bp (4); (b) Corynebacterium microorganisms, especially Corynebacterium, transformed with one or more (1), where these are replicative DNA; (c) production of branched-chain L-aa by fermentation of Corynebacterium bacteria in which the bnf and/or bnf genes (or equivalent sequences) are amplified, especially overexpressed; and (d) method for isolating the bnf and/or bnf genes. (1) is used for transformation of Corynebacterium bacteria being used for fermentative production of branched-chain amino acids, specifically leucine, isoleucine and valine, which are useful in medicine and animal nutrition. (1) can also be used as source of primers and probes for isolation of related sequences. Transformation with (1) increases yield of branched-chain amino acids. This sequence encodes the Corynebacterium glutamicum ATCC 14752 bnf and bnf proteins described in the method of the invention

Sequence 1271 BP; 265 A; 328 C; 318 G; 360 T; 0 U; 0 Other;

Query Match 100.0%; Score 1271; DB 4; Length 1271;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1271; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 GCGGATCATGATGATCTAGCTTCATATATATGCAAAATAGCCTTAGAGTGGCAAC 60
DB 1 GCGGATCATGATGATCTAGCTTCATATATATGCAAAATAGCCTTAGAGTGGCAAC 60
QY 61 TGGCAAAATATACCCGGCAATGTGTGATGATGATGATGATGATGATGATGATGAT 120
DB 61 TGGCAAAATATACCCGGCAATGTGTGATGATGATGATGATGATGATGATGATGAT 120
QY 61 TGGCAAAATATACCCGGCAATGTGTGATGATGATGATGATGATGATGATGATGAT 120
DB 61 TGGCAAAATATACCCGGCAATGTGTGATGATGATGATGATGATGATGATGATGAT 120
QY 121 TCATCAAGCTGAGAGTGTGCTCATCCAGGAGCCTTGGACAGATGATTAAGTTA 180
DB 121 TCATCAAGCTGAGAGTGTGCTCATCCAGGAGCCTTGGACAGATGATTAAGTTA 180
QY 121 TCATCAAGCTGAGAGTGTGCTCATCCAGGAGCCTTGGACAGATGATTAAGTTA 180
DB 121 TCATCAAGCTGAGAGTGTGCTCATCCAGGAGCCTTGGACAGATGATTAAGTTA 180
QY 181 TCGGCGCTACGAATTCGGGCAAGGTCTTAAACCTCCCTTGTGAGAGTTGGAGTGA 240
DB 181 TCGGCGCTACGAATTCGGGCAAGGTCTTAAACCTCCCTTGTGAGAGTTGGAGTGA 240
QY 181 TCGGCGCTACGAATTCGGGCAAGGTCTTAAACCTCCCTTGTGAGAGTTGGAGTGA 240
DB 181 TCGGCGCTACGAATTCGGGCAAGGTCTTAAACCTCCCTTGTGAGAGTTGGAGTGA 240
QY 241 CCCGATGGATTTGGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGT 300
DB 241 CCCGATGGATTTGGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGT 300
QY 301 CCCGATGGATTTGGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGT 360
DB 301 CCCGATGGATTTGGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGT 360
QY 361 TGTGGGCGAGGCGCCCTGGGCGCATTCGCTCACAACCATGCTGTGTAATTCGCGCA 420
DB 361 TGTGGGCGAGGCGCCCTGGGCGCATTCGCTCACAACCATGCTGTGTAATTCGCGCA 420
QY 421 CGTATTCATGCGTTTCAATTCGCGCTGATGTGTGTAATAAACCCTATGGCCGTTTCTA 480
DB 421 CGTATTCATGCGTTTCAATTCGCGCTGATGTGTGTAATAAACCCTATGGCCGTTTCTA 480
QY 481 TTGCGTTTTCGCGCTTATGAGAGAGCTACGAGTCACTGGGCGAGGCGCGAGCGTG 540
DB 481 TTGCGTTTTCGCGCTTATGAGAGAGCTACGAGTCACTGGGCGAGGCGCGAGCGTG 540
QY 481 TTGCGTTTTCGCGCTTATGAGAGAGCTACGAGTCACTGGGCGAGGCGCGAGCGTG 540
DB 481 TTGCGTTTTCGCGCTTATGAGAGAGCTACGAGTCACTGGGCGAGGCGCGAGCGTG 540
QY 541 GTGCGGCTGCGCATTCATTCATGCAATAGAGGTTTCACTCTACTGCGGTAATTCGCGG 600
DB 541 GTGCGGCTGCGCATTCATTCATGCAATAGAGGTTTCACTCTACTGCGGTAATTCGCGG 600
QY 601 TCTTACCGAGTGGCGATTCGCAAGTGAATTCCTTTTGAATTAAGGCTCGAGTTGCG 660
DB 601 TCTTACCGAGTGGCGATTCGCAAGTGAATTCCTTTTGAATTAAGGCTCGAGTTGCG 660
QY 661 CTTTGTCTCTCTTTTGTACAGCTGACTTGAATTCCTGCGGAGCAAAAAGAGATGCC 720
DB 661 CTTTGTCTCTCTTTTGTACAGCTGACTTGAATTCCTGCGGAGCAAAAAGAGATGCC 720

```

```

QY 721 TTCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 780
DB 721 TTCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 780
QY 781 CCTATTTGCGGCGCTGCTGATCTTCTGAGTGTGATGATGATGATGATGATGATGATGAT 840
DB 781 CCTATTTGCGGCGCTGCTGATCTTCTGAGTGTGATGATGATGATGATGATGATGATGAT 840
QY 841 AAAGCTCTAAATGACAACTGATTTCTCTGATATCTCTGTTGTCAGATGTCGA 900
DB 841 AAAGCTCTAAATGACAACTGATTTCTCTGATATCTCTGTTGTCAGATGTCGA 900
QY 901 GTCATTTACTTTGGCGCTCGCGCGGCTTCCGTTCTTAATCTTAAGCCCTACGTAATCA 960
DB 901 GTCATTTACTTTGGCGCTCGCGCGGCTTCCGTTCTTAATCTTAAGCCCTACGTAATCA 960
QY 961 CAATTTGTGGGCAAAATGAGCGATGTGATGTCAGCAGGAATCTTGCAATTTGACCGCA 1020
DB 961 CAATTTGTGGGCAAAATGAGCGATGTGATGTCAGCAGGAATCTTGCAATTTGACCGCA 1020
QY 1021 TCAAGCTTTCGCGAGCAATGCGATGATCTGAAAGACTTAACTTGTGCTAATGCGGCT 1080
DB 1021 TCAAGCTTTCGCGAGCAATGCGATGATCTGAAAGACTTAACTTGTGCTAATGCGGCT 1080
QY 1081 GCGATTTACGATGTGCGCGCATCTTCTTGGCGGTGCAACGACCTTGTGAGCGTGGCGCT 1140
DB 1081 GCGATTTACGATGTGCGCGCATCTTCTTGGCGGTGCAACGACCTTGTGAGCGTGGCGCT 1140
QY 1141 GCGACCATGCTTTTGTGAGCTGTGATCTTCTTCTTAAATCTGATTAATTAACAAAT 1200
DB 1141 GCGACCATGCTTTTGTGAGCTGTGATCTTCTTCTTAAATCTGATTAATTAACAAAT 1200
QY 1201 CCGCATGCCCTCAATTTGAAGGGAGTGGAGATTTTGAAGAACTAGAAAGGCTTAAG 1260
DB 1201 CCGCATGCCCTCAATTTGAAGGGAGTGGAGATTTTGAAGAACTAGAAAGGCTTAAG 1260
QY 1261 CAGACAGCGCT 1271
DB 1261 CAGACAGCGCT 1271

```

RESULT 2

AAH21112
ID AAH21112 standard; DNA; 1271 BP.

AAH21112;

05-SEP-2001 (first entry)

C. glutamicum DNA encoding bnf and bnfE.

L-amino acid production; bnf; bnfE; branched-chain amino acid; Corynebacterium; leucine; isoleucine; valine; medicine; animal nutrition; ds.

Corynebacterium glutamicum.

Key Location/Qualifiers
FH 101..856
FT /tag= "a"
FT /product= "bnf"
FT 853..1179
FT /tag= "b"
FT /product= "bnfE"

EP1096010-A1.

02-MAY-2001.

11-OCT-2000; 2000EP-00122057.

27-OCT-1999; 99DE-01051708.

XX 02-MAY-2001.
 XX 11-OCT-2000; 2000EP-00122057.
 XX 27-OCT-1999; 99DE-01051708.
 XX (DEGS) DEGUSA AG.
 XX (KERU) FORSCHUNGSZENTRUM JUELICH GMBH.
 XX Kennerknecht N, Eggeling L, Salm H, Pfeifferle W;
 XX WPI: 2001-391595/42.
 DR P-PSDB; AAB86247.
 XX New export genes from coryneform bacteria, useful for increasing
 PT fermentative production of branched-chain amino acids.
 XX
 PS Claim 5; Page 14-15; 23pp; German.
 XX This invention describes a novel isolated polynucleotide (I) containing
 CC at least one sequence that (i) is 70% identical with a sequence that
 CC encodes at least one of 251 amino acid (aa) (3) or 108 aa (5)
 CC polypeptides; (ii) encodes a polypeptide at least 70% identical with (3)
 CC or (5); (iii) is the complement of (i) or (ii), or (iv) contains at least
 CC 15 consecutive bases from (i)-(iii). The invention also describes (a)
 CC protein derived from sequences of 753 bp (2) or 324 bp (4); (b)
 CC coryneform microorganisms, especially Corynebacterium, transformed with
 CC one or more (1), where these are replicative DNA; (c) production of
 CC branched-chain L-aa by fermentation of coryneform bacteria in which the
 CC brnB and/or brnF genes (or equivalent sequences) are amplified,
 CC especially overexpressed; and (d) method for isolating the brnB and/or
 CC brnF genes. (1) is used for transformation of coryneform bacteria being
 CC used for fermentative production of branched-chain amino acids,
 CC specifically leucine, isoleucine and valine, which are useful in medicine
 CC and animal nutrition. (1) can also be used as source of primers and
 CC probes for isolation of related sequences. Transformation with (1)
 CC increases yield of branched-chain amino acids. This sequence encodes the
 CC Corynebacterium glutamicum ATCC 14752 brnF protein described in the
 CC method of the invention
 XX
 XX Sequence 753 BP; 140 A; 211 C; 194 G; 208 T; 0 U; 0 Other;
 SQ
 Query Match 59.2%; Score 753; DB 4; Length 753;
 Best Local Similarity 100.0%; Pred. No. 1.3e-230;
 Matches 753; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 361 AACCCATTGCCCGTTTCATTGCGTTTCGGGCTTATCGAAGAGCTACAGACTACT 420
 QY 521 GCGGCCAGGCGCCGACAGGCTGTGCGCGTGGCGACTTATCTCAATGCAATAGCGTTTAC 580
 DB 421 GCGGCCAGGCGCCGACAGGCTGTGCGCGTGGCGACTTATCTCAATGCAATAGCGTTTAC 480
 QY 581 TCGTACTGGGTATTCGCGCGGTCTCACCGGAGTGGCGATCGACAGTGTATTCCTTTGAA 640
 DB 481 TCGTACTGGGTATTCGCGCGGTCTCACCGGAGTGGCGATCGACAGTGTATTCCTTTGAA 540
 QY 641 ATTAGGCGCTCGAGTTCGCCCTTTGCTCTCTCTTTGTACGCTGACTTGGATTCTTCG 700
 DB 541 ATTAGGCGCTCGAGTTCGCCCTTTGCTCTCTCTTTGTACGCTGACTTGGATTCTTCG 600
 QY 701 CGAAGAAAAAGAGATCCCTTCTCGCTGCTCGAGGTTTGAAGCTTCAACATTGCTCTT 760
 DB 601 CGAAGAAAAAGAGATCCCTTCTCGCTGCTCGAGGTTTGAAGCTTCAACATTGCTCTT 660
 QY 761 GTGGTAATTCAGAGTCCGCTTATTTGCGCGCTGCTGATCTTCTTGGGCTCTGTGACC 820
 DB 661 GTGGTAATTCAGAGTCCGCTTATTTGCGCGCTGCTGATCTTCTTGGGCTCTGTGACC 720
 QY 821 ATCCGTAATCTTCTTCTTGGGAAAGGCTGTAA 853
 DB 721 ATCCGTAATCTTCTTCTTGGGAAAGGCTGTAA 753

RESULT 5
 ID AAB65254 standard; DNA; 753 BP.
 XX AAB65254;
 XX 26-SEP-2001 (first entry)
 DT C glutamicum coding sequence fragment SEQ ID NO: 289.
 XX
 XX Corynebacterium glutamicum.
 OS
 XX EPI108790-A2.
 XX
 XX 20-JUN-2001.
 PD 18-DEC-2000; 2000EP-00127688.
 XX
 XX 16-DEC-1999; 99JP-00377484.
 PR 07-APR-2000; 2000JP-00159162.
 PR 03-AUG-2000; 2000JP-00280988.
 XX
 XX (KYOW) KYOWA HAKKO KOGYO KK.
 PA
 PI Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;
 PI Tabeihi N, Senoh A, Ikeda M, Ozaki A;
 PI WPI: 2001-376931/40.
 DR P-PSDB; AAG90035.
 XX
 XX Novel polynucleotides derived from Coryneform bacteria, for identifying
 PT mutation point of a gene, measuring expression of a gene, analyzing
 PT expression profile or pattern of a gene and identifying homologous gene.
 XX
 PS Claim 8; SEQ ID NO 289; 246bp + Sequence Listing; English.
 XX
 XX The present invention provides a number of nucleotide and protein
 CC sequences from the Coryneform bacterium Corynebacterium glutamicum. These
 CC are useful for identifying the mutation point of a gene derived from a
 CC mutant of coryneform bacterium, measuring expression amount and analyzing
 CC the expression profile or expression pattern of a gene derived from
 CC Coryneform bacterium, and identifying a homologue of a gene derived from
 CC coryneform bacterium. Coryneform bacteria are useful for producing amino

CC acids, nucleic acids, vitamins, saccharides and organic acids.
 CC particularly L-lysine. The present sequence is a nucleic acid described
 CC in the exemplification of the invention. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from the European Patent Office
 XX

Sequence 753 BP; 140 A; 211 C; 194 G; 208 T; 0 U; 0 Other;

Query Match 59.2%; Score 753; DB 5; Length 753;

Best Local Similarity 100.0%; Pred. No. 1.3e-230; Indels 0; Gaps 0;

Matches 753; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

101 GTGCAAAAAGCAGAGATTCATTCAAGCCTGGAGGTGTGCGCATCCAGAGCCCTG 160
 1 GTGCAAAAAGCAGAGATTCATTCAAGCCTGGAGGTGTGCGCATCCAGAGCCCTG 60
 161 GAACCAAGATGATTAAGTTATCGCGCTACGAAATCGCGCAAGTCTTAAACCTCCCT 220
 61 GAACCAAGATGATTAAGTTATCGCGCTACGAAATCGCGCAAGTCTTAAACCTCCCT 120
 221 GCTGCAAGTTGGGAGATGACCGATGATGATGATGATGATGATGATGATGATGATGAT 280
 121 GCTGCAAGTTGGGAGATGACCGATGATGATGATGATGATGATGATGATGATGATGAT 180
 281 GAGTACGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 340
 181 GAGTACGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 240
 341 ATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 400
 241 ATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300
 401 TTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 460
 301 TTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 360
 461 AACCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 520
 361 AACCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420
 521 GGGGCGAGGCGCGAGGCTGTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTG 580
 421 GGGGCGAGGCGCGAGGCTGTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTG 480
 581 TCCCTAGTGGATATTCGGCGGCTTCAACCGAGTGGCGATCGCAGATGATTCCTTTTAA 640
 481 TCCCTAGTGGATATTCGGCGGCTTCAACCGAGTGGCGATCGCAGATGATTCCTTTTAA 540
 641 ATTAAGGAGGCTCGAGTTCGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 700
 541 ATTAAGGAGGCTCGAGTTCGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 600
 701 CGAAGCAAAAAGAGATTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 760
 601 CGAAGCAAAAAGAGATTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 660
 761 GTGGTAAATTCAGAGTTCAGAGGCTTATTTGGCGGCTGCTGATCTCTTGGGTCTGTGAC 820
 661 GTGGTAAATTCAGAGTTCAGAGGCTTATTTGGCGGCTGCTGATCTCTTGGGTCTGTGAC 720
 821 ATCCGATCT 883
 721 ATCCGATCT 753

RESULT 6
 ID ACA01969 standard; DNA; 564 BP.

ACA01969;
 04-JUN-2003 (first entry)

DE C. glutamicum derived ORF SEQ ID 1960.

XX Corynebacterium, nucleic acid array; fermentation; culture; ds.

XX Corynebacterium glutamicum.

XX DE10128510-A1.

XX 19-DEC-2002.

XX 13-JUN-2001; 2001DE-01028510.

XX 13-JUN-2001; 2001DE-01028510.

XX (DEGS) DEGUSA AG.

XX Farwick M, Moekel B, Pfefferle W, Bathe B, Huthmacher K;

XX WPI; 2003-279970/28.

XX New nucleic acid array useful for monitoring mRNA expression of
 PT Corynebacterium glutamicum during fermentation, comprising nucleic acid
 PT from Corynebacterium glutamicum.

XX Claim 1; Page 647; 709pp; German.

XX This invention describes a novel nucleic acid array involving
 CC Corynebacterium glutamicum polynucleotides. The arrays are used to
 CC analyze C. glutamicum, particularly for monitoring a fermentation process
 CC to determine expression levels of C. glutamicum cellular mRNA. Such
 CC monitoring particularly differentiates between expression levels of
 CC different strains of C. glutamicum and allows the adjustment of different
 CC culture and fermentation conditions. ACA00010-ACA02188 represent C.
 CC glutamicum derived polynucleotides described in the disclosure of the
 CC invention

XX Sequence 564 BP; 91 A; 167 C; 144 G; 162 T; 0 U; 0 Other;

Query Match 43.5%; Score 553; DB 7; Length 564;

Best Local Similarity 99.8%; Pred. No. 1.8e-166; Indels 1; Gaps 1;

Matches 564; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

289 ATGATGGGAGCCCGACCTGTTTCCGCGCTGATTTTCGGGCTGACCGGAAATGCTGAT 348
 1 ATGATGGGAGCCCGACCTGTTTCCGCGCTGATTTTCGGGCTGACCGGAAATGCTGAT 60
 349 CATGCCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 408
 61 CATGCCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 119
 409 GAACTTCGCGACGATATTCATGCGGCTGATTCGCGCTGATGATGATGATGATGATGAT 468
 120 GAACTTCGCGACGATATTCATGCGGCTGATTCGCGCTGATGATGATGATGATGATGAT 179
 469 TGCCCGTTTATTCGTTTTCGCGCTTATTCGCGCTTATTCGCGCTTATTCGCGCTTAT 528
 180 TGCCCGTTTATTCGTTTTCGCGCTTATTCGCGCTTATTCGCGCTTATTCGCGCTTAT 239
 529 GCCCGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 588
 240 GCCCGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 299
 589 GGTATTCGCGGCTTACCGGAGTGGGATCGGAGGATGATGATGATGATGATGATGATGAT 648
 300 GGTATTCGCGGCTTACCGGAGTGGGATCGGAGGATGATGATGATGATGATGATGATGAT 359
 649 CCTGAGTTCGCGCTTCT 708
 360 CCTGAGTTCGCGCTTCT 419
 709 AAAGGAGATTCCT 768
 420 AAAGGAGATTCCT 479

```

QY 769 TCCAGGTGACGCGCTATTTGCGGCGTGCATCTTCTGGGTCTGTGACCATCCGGA 828
DB 480 TCCAGGTGACGCGCTATTTGCGGCGTGCATCTTCTGGGTCTGTGACCATCCGGA 539
QY 829 CTTCTTCTGGGAAAGGCTGCTAA 853
DB 540 CTTCTTCTGGGAAAGGCTGCTAA 564

RESULT 7
AAf61693/c
ID AAF61693 standard; DNA; 397 BP.
AC AAF61693;
XX
XX 12-JUL-2001 (first entry)
DE C. glutamicum lrp encoding DNA fragment SEQ ID 7.
KW Fermentation; L-amino acid production; lrp gene; lysine; isoleucine;
KM medicine; animal feed supplement; de.
XX Corynebacterium glutamicum.
XX EP1090993-A1.
XX 11-APR-2001.
XX 29-SEP-2000; 2000EP-00121159.
XX 05-OCT-1999; 99DE-01047792.
XX (DEGS ) DEGUSA-HUELS AG.
XX Moockel B, Pfeifferle W, Puehler A, Kalinowski J, Bathe B;
XX WPI; 2001-292927/31.
XX New lrp gene from coryneform bacteria, used to prepare transformants with
XX increased synthesis of amino acids, particularly lysine and isoleucine.
XX Example 3; Page 16; 22pp; German.
XX This invention describes a novel isolated nucleic acid (I) from
XX coryneform bacteria used for the fermentative production of selected L-
XX amino acids, by fermenting the amino acid-producing coryneform in which
XX at least the lrp gene has been weakened or amplified, then isolating
XX amino acids that have accumulated in the medium or cells. (I) is used to
XX transform coryneforms for production of L-amino acids, specifically
XX lysine and isoleucine, which are used in medicine and particularly as
XX animal feed supplement. It may also be used as probes and primers for
XX isolating related sequences. Regulating expression of (I) improves
XX production of amino acids, especially of L-lysine. This sequence encodes
XX a fragment of the Corynebacterium glutamicum lrp protein which is used in
XX the method described in the invention
SQ Sequence 397 BP; 96 A; 100 C; 104 G; 97 T; 0 U; 0 Other;
Query Match 28.8%; Score 365.6; DB 4; Length 397;
Best Local Similarity 98.9%; Pred. No. 2.2e-106;
Matches 368; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 9 AATGAATCTAGCTCATATATTTGACAAATAGCCTAGTTAGTGGCGGAACTGCAACA 68
DB 376 AATGAATCTAGCTCATATATTTGACAAATAGCCTAGTTAGTGGCGGAACTGCAACA 317
QY 69 AATGAATCTAGCTCATATATTTGACAAATAGCCTAGTTAGTGGCGGAACTGCAACA 128
DB 316 AATGAATCTAGCTCATATATTTGACAAATAGCCTAGTTAGTGGCGGAACTGCAACA 257
QY 129 GCGTGGAGGTGTCGCATCCAAAGGCGCGTGGAAAGCATGATTAAGTTATCGCGCCT 188

```

```

DB 256 GCGTGGAGGTGTCGCATCCAAAGGCGCGTGGAAAGCATGATTAAGTTATCGCGCCT 197
QY 189 ACGAATAGCGGGAAGTGTAAAAACCTCCCTGTCGACGATTTGGGACATGATACCGATTG 248
DB 196 ACGAATAGCGGGAAGTGTAAAAACCTCCCTGTCGACGATTTGGGACATGATACCGATTG 137
QY 249 GATTAAGCTTGTGCTCTGTTGTTATTCATACGCTGCAATGTGTGGGACAGCCCATGT 308
DB 136 GATTAAGCTTGTGCTCTGTTGTTATTCATACGCTGCAATGTGTGGGACAGCCCATGT 77
QY 309 TTTCCGCGCTGATTTTTCGCGGAGCTCCACCGAAATCTGTATGCGCTCGTTGTGGCG 368
DB 76 TTTCCGCGCTGATTTTTCGCGGAGCTCCACCGAAATCTGTATGCGCTCGTTGTGGCG 17
QY 369 CAGCGCCCTCG 380
DB 16 CAGCGCCCTCG 5

RESULT 8
AAf61695/c
ID AAF61695 standard; DNA; 778 BP.
AC AAF61695;
XX
XX 12-JUL-2001 (first entry)
DE C. glutamicum lrp encoding DNA fragment SEQ ID 9.
KW Fermentation; L-amino acid production; lrp gene; lysine; isoleucine;
KM medicine; animal feed supplement; ds.
XX Corynebacterium glutamicum.
XX EP1090993-A1.
XX 11-APR-2001.
XX 29-SEP-2000; 2000EP-00121159.
XX 05-OCT-1999; 99DE-01047792.
XX (DEGS ) DEGUSA-HUELS AG.
XX Moockel B, Pfeifferle W, Puehler A, Kalinowski J, Bathe B;
XX WPI; 2001-292927/31.
XX New lrp gene from coryneform bacteria, used to prepare transformants with
XX increased synthesis of amino acids, particularly lysine and isoleucine.
XX Example 3; Page 17; 22pp; German.
XX This invention describes a novel isolated nucleic acid (I) from
XX coryneform bacteria used for the fermentative production of selected L-
XX amino acids, by fermenting the amino acid-producing coryneform in which
XX at least the lrp gene has been weakened or amplified, then isolating
XX amino acids that have accumulated in the medium or cells. (I) is used to
XX transform coryneforms for production of L-amino acids, specifically
XX lysine and isoleucine, which are used in medicine and particularly as
XX animal feed supplement. It may also be used as probes and primers for
XX isolating related sequences. Regulating expression of (I) improves
XX production of amino acids, especially of L-lysine. This sequence encodes
XX a fragment of the Corynebacterium glutamicum lrp protein which is used in
XX the method described in the invention
SQ Sequence 778 BP; 175 A; 208 C; 207 G; 186 T; 0 U; 0 Other;
Query Match 28.8%; Score 365.6; DB 4; Length 778;
Best Local Similarity 98.9%; Pred. No. 3.2e-106;
Matches 368; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 9 AATGAATCTAGCTCATATATTTGACAAATAGCCTAGTTAGTGGCGGAACTGCAACA 68

```

```

Cc |||||
Db 376 AATGAATCTAGCTCATATATATGCAATAGCCTAGTGGTGGCGCAAACTGGCAACA 317
Qy 69 AAACATACCCGGCAATGTGTGTATGTATGTAGTGTGCAAAAAACCAAGATTCATTCAA 128
Db 316 AAACATACCCGGCAATGTGTGTATGTATGTAGTGTGCAAAAAACCAAGATTCATTCAA 257
Qy 129 GCTTGAAGGTGTGCGCATCCAGAGCGCTGGAACCAAGATTAAGTTATGCGCGCT 188
Db 256 GCTTGAAGGTGTGCGCATCCAGAGCGCTGGAACCAAGATTAAGTTATGCGCGCT 197
Qy 169 ACGAATATGCGGCAAGGTCTTAAAAACCTCTTGTGCGAGGTGGCATGTACCCGATTG 248
Db 196 ACGAATATGCGGCAAGGTCTTAAAAACCTCTTGTGCGAGGTGGCATGTACCCGATTG 137
Qy 249 GATTGCGTTTGTGTCTTGTATTATCAATACGCTACGAATGTGTGGCAGCCCACTGT 308
Db 136 GATTGCGTTTGTGTCTTGTATTATCAATACGCTACGAATGTGTGGCAGCCCACTGT 77
Qy 309 TTTCCGCGCTGATTTTTCGCGGCGCTCAGCGAATGCTGTATGCGCTCGTTGTGGCG 368
Db 76 TTTCCGCGCTGATTTTTCGCGGCGCTCAGCGAATGCTGTATGCGCTCGTTGTGGCG 17
Qy 369 CAGCGCCCTG 380
Db 16 CAGCGCCCTG 5

RESULT 9
AAH21111
ID AAH21111 standard; DNA; 324 BP.
XX
AC AAH21111;
DT 05-SEP-2001 (first entry)
DE C. glutamicum b7m DNA.
XX
KM L-amino acid production; b7m; b7m; branched-chain amino acid;
KM coryneform bacterium; leucine; isoleucine; valine; medicine;
KM animal nutrition; ds.
XX
OS Corynebacterium glutamicum.
XX
PN EP1096010-A1.
XX
PD 02-MAY-2001.
XX
PF 11-OCT-2000; 2000EP-00122057.
XX
PR 27-OCT-1999; 99DE-01051708.
XX
PA (DEBS ) DEGUSA AG.
PA (KERJ ) FORSCHUNGSZENTRUM JUELICH GMBH.
PI Kennerknecht N, Eggeling L, Sahn H, Pfeifferle W;
XX
XX WPI; 2001-391595/42.
DR P-PSDB; AAB86248.
XX
XX New export genes from coryneform bacteria, useful for increasing
PT fermentative production of branched-chain amino acids.
XX
XX Claim 5; Page 16; 23pp; German.
XX
XX This invention describes a novel isolated polynucleotide (I) containing
CC at least one sequence that (i) is 70% identical with a sequence that
CC encodes at least one of 251 amino acid (aa) (3) or 108 aa (5)
CC polypeptides; (ii) encodes a polypeptide at least 70% identical with (3)
CC or (5); (iii) is the complement of (i) or (ii), or (iv) contains at least
CC 15 consecutive bases from (i)-(iii). The invention also describes (a)
CC protein derived from sequences of 753 bp (2) or 324 bp (4); (b)
CC coryneform microorganisms, especially Corynebacterium, transformed with

```

```

Cc one or more (I), where these are replicative DNA; (c) production of
Cc branched-chain L-a by fermentation of coryneform bacteria in which the
Cc b7m and/or b7m genes (or equivalent sequences) are amplified,
Cc especially overexpressed; and (d) method for isolating the b7m and/or
Cc b7m genes. (I) is used for transformation of coryneform bacteria being
Cc used for fermentative production of branched-chain amino acids,
Cc specifically leucine, isoleucine and valine, which are useful in medicine
Cc and animal nutrition. (I) can also be used as source of primers and
Cc probes for isolation of related sequences. Transformation with (I)
Cc increases yield of branched-chain amino acids. This sequence encodes the
Cc coryneform glutamicum ATCC 14752 b7m protein described in the
Cc method of the invention
Cc
Cc Sequence 324 BP; 60 A; 79 C; 80 G; 105 T; 0 U; 0 Other;
Cc
Cc Query Match 25.5%; Score 324; DB 4; Length 324;
Cc Best Local Similarity 100.0%; Pred. No. 4.4e-93;
Cc Matches 324; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 853 ATGACAACTGATTTCTCCCTGATTTCCCTTGTGTGTGCGAGTATGCGACTTACTTT 912
Db 1 ATGACAACTGATTTCTCCCTGATTTCTCCCTTGTGTGTGCGAGTATGCGACTTACTTT 60
Qy 913 GCGCTCCGCGCGGCTTCGCTTCTTAATCCTTAAGCCCTTACGTAATCACAATTGTGGGC 972
Db 61 GCGCTCCGCGCGGCTTCGCTTCTTAATCCTTAAGCCCTTACGTAATCACAATTGTGGGC 120
Qy 973 AAATAGCGATGTGATGCGAGAGAGATCCCTTGTGATTTTACCGCATCAAGCTTGGC 1032
Db 121 AAATAGCGATGTGATGCGAGAGAGATCCCTTGTGATTTTACCGCATCAAGCTTGGC 180
Qy 1033 AGCAATGCGATGATCTGAAGAAGCTTAACCTTGTGATTTTACCGCATCAAGCTTGGC 1092
Db 181 AGCAATGCGATGATCTGAAGAAGCTTAACCTTGTGATTTTACCGCATCAAGCTTGGC 240
Qy 1093 GTGGCGCATTTTGTGGCGGTCGAGCGACCTTGTGAGGCTTGGCGCTGGCATTCGTT 1152
Db 241 GTGGCGCATTTTGTGGCGGTCGAGCGACCTTGTGAGGCTTGGCGCTGGCATTCGTT 300
Qy 1153 TTTGTGAGCTGTGATCTTTTC 1176
Db 301 TTTGTGAGCTGTGATCTTTTC 324

RESULT 10
AAH65255
ID AAH65255 standard; DNA; 324 BP.
XX
AC AAH65255;
DT 26-SEP-2001 (first entry)
DE C glutamicum coding sequence fragment SEQ ID NO: 250.
XX
KM Coryneform bacterium; amino acid synthesis; vitamin; saccharide;
KM organic acid synthesis; ds.
XX
OS Corynebacterium glutamicum.
XX
PN EP108790-A2.
XX
PD 20-JUN-2001.
XX
PF 18-DEC-2000; 2000EP-00127688.
XX
PR 16-DEC-1999; 99JP-00377484.
PR 07-APR-2000; 2000JP-00159162.
PR 03-AUG-2000; 2000JP-00280988.
XX
XX (KYOW ) KYOWA HAKKO KOGYO KK.
XX
XX Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;
PI Tateishi N, Semoh A, Ikeda M, Ozaki A;

```



```

FT      -10_signal      /*tag= a
FT      88.93
FT      /*tag= b
FT      CDS      151.615
FT      /*tag= c
FT      /product= "lrp"
XX
XX      EP1090993-A1.
XX
XX      11-APR-2001:
XX
XX      29-SEP-2000; 2000EP-00121159.
XX
XX      05-OCT-1999; 99DE-01047792.
XX
XX      (DEGS ) DEGUSA-HUELS AG.
XX
XX      Moeckel B, Pfefferle W, Puhler A, Kalinowski J, Bathe B;
XX      WPI; 2001-292927/31.
XX      P-PsDB; AAB70881.
XX
XX      New lrp gene from coryneform bacteria, used to prepare transformants with
XX      increased synthesis of amino acids, particularly lysine and isoleucine.
XX
XX      Claim 4; Page 14-15; 22pp; German.
XX
XX      This invention describes a novel isolated nucleic acid (I) from
XX      coryneform bacteria used for the fermentative production of selected L-
XX      amino acids, by fermenting the amino acid-producing coryneform in which
XX      at least the lrp gene has been weakened or amplified, then isolating
XX      amino acids that have accumulated in the medium or cells. (I) is used to
XX      transform coryneforms for production of L-amino acids, specifically
XX      lysine and isoleucine, which are used in medicine and particularly as
XX      animal feed supplement. It may also be used as probes and primers for
XX      isolating related sequences. Regulating expression of (I) improves
XX      production of amino acids, especially of L-lysine. This sequence encodes
XX      the Corynebacterium glutamicum lrp protein which is used in the method
XX      described in the invention
XX
XX      Sequence 715 BP; 166 A; 173 C; 178 G; 198 T; 0 U; 0 Other;
XX
XX      Query Match      14.6%; Score 185; DB 4; Length 715;
XX      Best Local Similarity 100.0%; Pred. No. 2.9e-48;
XX      Matches 185; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX      QY      1 GCGCGATCATGGAATCTAGCTTCATATATTGCAACAATAGCCTAGTTGAGTGGCCAAAC 60
XX      DB      185 GCGCGATCATGGAATCTAGCTTCATATATTGCAACAATAGCCTAGTTGAGTGGCCAAAC 126
XX      QY      61 TGGCAACAAACTACCCGGCAATTGTGATGATTGTAGTGGCAAAAAACCGAAGAT 120
XX      DB      125 TGGCAACAAACTACCCGGCAATTGTGATGATTGTAGTGGCAAAAAACCGAAGAT 66
XX      QY      121 TCATTCAAGCCTGAGAGTGTGCGCATCCAGGACAGCCCTGGAACCGATGATTAAGTTA 180
XX      DB      65 TCATTCAAGCCTGAGAGTGTGCGCATCCAGGACAGCCCTGGAACCGATGATTAAGTTA 6
XX      QY      181 TCGGC 185
XX      DB      5 TCGGC 1
XX
XX      RESULT 13
XX      ID      ACA01968 standard; DNA; 177 BP.
XX      AC      ACA01968;
XX      XX
XX      DT      04-JUN-2003 (first entry)
XX      XX
XX      DE      C. glutamicum derived ORF SEQ ID 1959.
XX

```

```

XX      Coryneform; nucleic acid array; fermentation; culture; ds.
XX      Corynebacterium glutamicum.
XX      DE10128510-A1.
XX      19-DEC-2002.
XX      13-JUN-2001; 2001DE-01028510.
XX      PF      13-JUN-2001; 2001DE-01028510.
XX      PR      13-JUN-2001; 2001DE-01028510.
XX      XX
XX      (DEGS ) DEGUSA AG.
XX
XX      Farwick M, Moeckel B, Pfefferle W, Bathe B, Hutmacher K;
XX      WPI; 2003-279970/28.
XX
XX      New nucleic acid array useful for monitoring mRNA expression of
XX      Corynebacterium glutamicum during fermentation, comprising nucleic acid
XX      from Corynebacterium glutamicum.
XX
XX      Claim 1; Page 647; 709pp; German.
XX
XX      This invention describes a novel nucleic acid array involving
XX      Corynebacterium glutamicum polynucleotides. The arrays are used to
XX      analyse C. glutamicum, particularly for monitoring a fermentation process
XX      to determine expression levels of C. glutamicum cellular mRNA. Such
XX      monitoring particularly differentiates between expression levels of
XX      different strains of C. glutamicum and allows the adjustment of different
XX      culture and fermentation conditions. ACA00010-ACA02188 represent C.
XX      glutamicum derived polynucleotides described in the disclosure of the
XX      invention
XX
XX      Sequence 177 BP; 54 A; 40 C; 47 G; 36 T; 0 U; 0 Other;
XX
XX      Query Match      13.9%; Score 177; DB 7; Length 177;
XX      Best Local Similarity 100.0%; Pred. No. 4.8e-46;
XX      Matches 177; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX      QY      30 TTGCACATAGCCTAGTTGAGTGGCGCAACTGGCAACAACTACCCGGCAATTGTGTG 89
XX      DB      1 TTGCACATAGCCTAGTTGAGTGGCGCAACTGGCAACAACTACCCGGCAATTGTGTG 60
XX      QY      90 ATGATTGTAGTGGCAAAAAACCGAAGATTTATTCAAGCCTGGAGGTGCCCATCCA 149
XX      DB      61 ATGATTGTAGTGGCAAAAAACCGAAGATTTATTCAAGCCTGGAGGTGCCCATCCA 120
XX      QY      150 AGGAGCCCTGGAAACGATGATTAAGGTTATCGCGCTACGAAATCGCCCAAGTTC 206
XX      DB      121 AGGAGCCCTGGAAACGATGATTAAGGTTATCGCGCTACGAAATCGCCCAAGTTC 177
XX
XX      RESULT 14
XX      ID      AAH68373
XX      AAH68373 standard; DNA; 711 BP.
XX      AC      AAH68373;
XX      XX
XX      DT      26-SEP-2001 (first entry)
XX      XX
XX      DE      C glutamicum coding sequence fragment SEQ ID NO: 3408.
XX      XX
XX      Coryneform bacterium; amino acid synthesis; vitamin; saccharide;
XX      organic acid synthesis; ds.
XX      Corynebacterium glutamicum.
XX      OS
XX      PN      EP108790-A2.
XX      PD      20-JUN-2001.
XX      PF      18-DEC-2000; 2000EP-00127688.
XX

```



```
Db 233 TCATGATGATTTCCGCCGACATTTTCTAAGTCTCACCTTCCACGCCACCGCATCAAGT 292
Qy 462 ACCCCATTGCCCGTTTCTATTTGCTTTGCGCTTATGACGAGCCTTACGAGTCACTG 521
Db 293 CCGGGCGCCCGCCGCTTATTCACCTTACCGAGCTTACCGAGCTTACCGCATCGTGT 352
Qy 522 CCGCCAGGCGCCGACGCTGTGCGCTG--GCGACTTATCTCAATGCAATAGCGTTTC 578
Db 353 CAGCCCGCCCACTTGCGCATATCAGTGGACGCGGGTGTACCGTTCAATTTGTGCC 412
Qy 579 ACTCTACTGGGTATTCGCGGTCTCACCGAGTGGCGATGCGAGATTGATCTTTTG 638
Db 413 AAGCTCTGTGGGTATCCGAGGATATATGCGCCTTGTGTGTCAAGTGTGCCGATG 472
Qy 639 AAATTAAAGGCTCGAGTTGCGCCTTGTCTCTCTTTGTCAAGCTGACTTGAATTCCT 698
Db 473 ATCTAAAGGCATGATTTTGCCCTGACCGCCTGTGTGTGCTGGCGTGGAGGCAT 532
Qy 699 GCCGAAAGGAAAGCAGATCCCTTCTGTGCTGCTGCGAGTTGAGCTTCAACATTTGCTC 758
Db 533 TCAGAAATTAACAAGATTAATCGTGCATTAATCGCGGTGATTTGCTGTGTTCCG 592
Qy 759 TTGTGTATTTCAAGTCAAGGCTTATTTGCGCGCTGCTGATCTTTGGGTCTGTGA 818
Db 593 GTTTGTGGCGCCGAGCAGATGTGTATGCTTTGACACGTACTTTTGAATCTTC 652
Qy 819 CCATCGGTACTTCTTC 835
Db 653 TTCTCGCGCTCGCTTC 669
```

Search completed: April 19, 2004, 11:41:56
Job time: 595.9 secs

Db 181 TCGGCGCTACGAAATCGCGCAGAGTCTAAAACTCCCTTGTGCGAGTTGGGCAATGTA 240
 QY 241 CCCGATGTATATGGGTTGGTCTCTGGTTATTCATATACGGCTACGAATGATGAGGAGC 300
 Db 241 CCCGATGTATATGGGTTGGTCTCTGGTTATTCATATACGGCTACGAATGATGAGGAGC 300
 QY 301 CCCACTGTTTTCCGGCTGATATTTCCGGGCTCCACCGAAATGCTGATCGCCCTGCT 360
 Db 301 CCCACTGTTTTCCGGGCTGATATTTCCGGGCTCCACCGAAATGCTGATCGCCCTGCT 360
 QY 361 TGTGGGCGAGAGCCCTTGGGCGCATCGCGCTCACCAATGCTGTGAATCTCCGCA 420
 Db 361 TGTGGGCGAGAGCCCTTGGGCGCATCGCGCTCACCAATGCTGTGAATCTCCGCA 420
 QY 421 CGTATCTATGAGGTTTCAATCCCGCTGCATGTGTCAAAAACCCCATTTGCCGTTCTA 480
 Db 421 CGTATCTATGAGGTTTCAATCCCGCTGCATGTGTCAAAAACCCCATTTGCCGTTCTA 480
 QY 481 TTGGGTTTCCGGCTTATGAGAGAGCTACGAGTCTGCGGCGAGCCCGCAGGCTG 540
 Db 481 TTGGGTTTCCGGCTTATGAGAGAGCTACGAGTCTGCGGCGAGCCCGCAGGCTG 540
 QY 541 GTGCGGCTGAGCACTTATCTCAATGCAATAGCGTTTCACTCTACTGAGGTTTGGCGG 600
 Db 541 GTGCGGCTGAGCACTTATCTCAATGCAATAGCGTTTCACTCTACTGAGGTTTGGCGG 600
 QY 601 TCTCACCGAGATGCGCATGCGAGAGTTGATCTCTTTGAAATTAAGGCGCTGATGCG 660
 Db 601 TCTCACCGAGATGCGCATGCGAGAGTTGATCTCTTTGAAATTAAGGCGCTGATGCG 660
 QY 661 CCTTGTCTCTCTTTGTGTACGCTGACTTTGAGATTCCTCCGAGAGAAAGACAGATCC 720
 Db 661 CCTTGTCTCTCTTTGTGTACGCTGACTTTGAGATTCCTCCGAGAGAAAGACAGATCC 720
 QY 721 TTCTGTCTGTCTGAGAGTTTGAAGTTCACCAATGCTTGTGTATTCAGTTCAGGCT 780
 Db 721 TTCTGTCTGTCTGAGAGTTTGAAGTTCACCAATGCTTGTGTATTCAGTTCAGGCT 780
 QY 781 CCTATTGGGCGGCTGCTGATCTCTTGGGCTGTTGAGCAATCCGGTACTTCTTCTGGG 840
 Db 781 CCTATTGGGCGGCTGCTGATCTCTTGGGCTGTTGAGCAATCCGGTACTTCTTCTGGG 840
 QY 841 AAAAGCTGTAAATGACAACTGATTTCTCTGTATTTCTCTTGTGTGCGAGTATGCA 900
 Db 841 AAAAGCTGTAAATGACAACTGATTTCTCTGTATTTCTCTTGTGTGCGAGTATGCA 900
 QY 901 GTGATTAATTTTGGGCTCCGGGCGGTTCCGTTCTTAATGCTTAAGCCCTTACGTATCA 960
 Db 901 GTGATTAATTTTGGGCTCCGGGCGGTTCCGTTCTTAATGCTTAAGCCCTTACGTATCA 960
 QY 961 CAATTTGTGGGCAAAATGCGAGATGTGATGCCAGAGAACTCTTGGCAATTTTGAACGCA 1020
 Db 961 CAATTTGTGGGCAAAATGCGAGATGTGATGCCAGAGAACTCTTGGCAATTTTGAACGCA 1020
 QY 1021 TCAAGCTTTGCGAGCAATGCGATAGATCTGAAGACTTAACTTTGCTCATTTGCCGTT 1080
 Db 1021 TCAAGCTTTGCGAGCAATGCGATAGATCTGAAGACTTAACTTTGCTCATTTGCCGTT 1080
 QY 1081 GCGATTAAGTGTGCGGCTCTTCTTGGGCTGAGAGCACTTGTGAGCGTTGGCGCT 1140
 Db 1081 GCGATTAAGTGTGCGGCTCTTCTTGGGCTGAGAGCACTTGTGAGCGTTGGCGCT 1140
 QY 1141 GCGACCATGCTTTTGTGAGCTGTGATCTTTTCTAAATCTGCTAAATAAACAATAAT 1200
 Db 1141 GCGACCATGCTTTTGTGAGCTGTGATCTTTTCTAAATCTGCTAAATAAACAATAAT 1200
 QY 1201 CCGATGCTCTCAATTTGAAGGGATGCGGATTTTGAAGGAACCTAGAAAAGGCTTAAG 1260
 Db 1201 CCGATGCTCTCAATTTGAAGGGATGCGGATTTTGAAGGAACCTAGAAAAGGCTTAAG 1260
 QY 1261 CAGACAGCGCT 1271
 Db 1261 CAGACAGCGCT 1271

Db 1261 CAGACAGCGCT 1271
 RESULT 2
 US-09-471-803A-6
 ; Sequence 6, Application US/09471803A
 ; Patent No. 6613545
 ; GENERAL INFORMATION:
 ; APPLICANT: KENNERNKCHT, NICOLE
 ; APPLICANT: SAHM, HERMANN
 ; APPLICANT: EGELING, LOTAR
 ; APPLICANT: PEPERLE, WALTER
 ; TITLE OF INVENTION: NUCLEOTIDE SEQUENCE CODING FOR THE EXPORT OF
 ; TITLE OF INVENTION: BRANCHED CHAIN AMINO ACIDS, PROCESS FOR THE
 ; FILE REFERENCE: 21123/265496/MAS
 ; CURRENT FILING DATE: 1999-12-23
 ; PRIOR APPLICATION NUMBER: US/09/471, 803A
 ; PRIOR FILING DATE: 1999-10-27
 ; NUMBER OF SEQ ID NOS: 12
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 6
 ; LENGTH: 1271
 ; TYPE: DNA
 ; ORGANISM: Corynebacterium glutamicum
 ; FEATURE:
 ; NAME/KEY: gene
 ; LOCATION: (101)..(853)
 ; OTHER INFORMATION: brnF
 ; NAME/KEY: gene
 ; LOCATION: (853)..(1176)
 ; OTHER INFORMATION: brnE
 ; OTHER INFORMATION: ANCC13032
 US-09-471-803A-6
 Query Match 100.0%; Score 1271; DB 4; Length 1271;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1271; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GCGGATCATGGAATGATCTAGCTCATATATATGACAAATAGCTAGTGAAGTCCGAAC 60
 Db 1 GCGGATCATGGAATGATCTAGCTCATATATATGACAAATAGCTAGTGAAGTCCGAAC 60
 QY 61 TGGCAACAAACTACCCGGCAATTTGTGTGATGATGTAGTGCAGAAACGCAAGAT 120
 Db 61 TGGCAACAAACTACCCGGCAATTTGTGTGATGATGTAGTGCAGAAACGCAAGAT 120
 QY 121 TCATTCAAGCTGAGAGTGTCCGATCCAGAGCAAGCCCTGGAAACGAGATGAAGGTTA 180
 Db 121 TCATTCAAGCTGAGAGTGTCCGATCCAGAGCAAGCCCTGGAAACGAGATGAAGGTTA 180
 QY 181 TCGGCGCTACGAAATCGCGCAAGGCTTAAAACTCCCTTGTGTCAGGTTTGGGCAATGTA 240
 Db 181 TCGGCGCTACGAAATCGCGCAAGGCTTAAAACTCCCTTGTGTCAGGTTTGGGCAATGTA 240
 QY 241 CCCGATGTATATGGGTTGGTCTCTGGTTATTCATATACGGCTACGAATGATGAGGAGC 300
 Db 241 CCCGATGTATATGGGTTGGTCTCTGGTTATTCATATACGGCTACGAATGATGAGGAGC 300
 QY 301 CCCACTGTTTTCCGGCTGATATTTCCGGGCTCCACCGAAATGCTGATCGCCCTGCT 360
 Db 301 CCCACTGTTTTCCGGCTGATATTTCCGGGCTCCACCGAAATGCTGATCGCCCTGCT 360
 QY 361 TGTGGGCGAGAGCCCTTGGGCGCATCGCGCTCACCAATGCTGTGAATCTCCGCA 420
 Db 361 TGTGGGCGAGAGCCCTTGGGCGCATCGCGCTCACCAATGCTGTGAATCTCCGCA 420
 QY 421 CGTATCTATGAGGTTTCAATCCCGCTGCATGTGTCAAAAACCCCATTTGCCGTTCTA 480
 Db 421 CGTATCTATGAGGTTTCAATCCCGCTGCATGTGTCAAAAACCCCATTTGCCGTTCTA 480
 QY 481 TTGGGTTTCCGGCTTATGAGAGAGCTACGAGTCTGCGGCGAGCCCGCAGGCTG 540
 Db 481 TTGGGTTTCCGGCTTATGAGAGAGCTACGAGTCTGCGGCGAGCCCGCAGGCTG 540

```

Db      481  TTGCGTTTGGCGCTTTCAGCAGAGCCTAGCGATCGCGCCGAGCGCTG 540
Qy      541  GTGCGGCTGGCGCATTAATCTCAATGCAAAATAGCGTTTCACTCTACTAGGTAATTCGCGG 600
Db      541  GTGCGGCTGGCGCATTAATCTCAATGCAAAATAGCGTTTCACTCTACTAGGTAATTCGCGG 600
Qy      601  TCTCACCGGAGTGGCGATCGCAGAGTTGATCTCTTTGAAATTAGGGCTCGAGTTGCG 660
Db      601  TCTCACCGGAGTGGCGATCGCAGAGTTGATCTCTTTGAAATTAGGGCTCGAGTTGCG 660
Qy      661  CCTTGGCTCTCTCTTTGCAAGCTGATCTTGGATTCCTGCGGAGCAAGAAAGAGATCCC 720
Db      661  CCTTGGCTCTCTCTTTGCAAGCTGATCTTGGATTCCTGCGGAGCAAGAAAGAGATCCC 720
Qy      721  TTCTGTGCTGCTGCGAGGTTTGAAGCTTCAACATTCCTTTGTTGTAATTCAGGTGAGC 780
Db      721  TTCTGTGCTGCTGCGAGGTTTGAAGCTTCAACATTCCTTTGTTGTAATTCAGGTGAGC 780
Qy      781  CCTATTGCGGCGCTGCTGATCTTCTTGGGTCGTGACCATCGGTAATCTTCTTGGG 840
Db      781  CCTATTGCGGCGCTGCTGATCTTCTTGGGTCGTGACCATCGGTAATCTTCTTGGG 840
Qy      841  AAAGGCTGCTAAATGACAACTGATTTCTCTGTAATCTCTCTTGTGTGCGAGTAATGCA 900
Db      841  AAAGGCTGCTAAATGACAACTGATTTCTCTGTAATCTCTCTTGTGTGCGAGTAATGCA 900
Qy      901  GTCAATTAATTTGGCGCTTCGCGCGGTTCCGTTCTTAATCTTAAAGCCCTACGTGAAATCA 960
Db      901  GTCAATTAATTTGGCGCTTCGCGCGGTTCCGTTCTTAATCTTAAAGCCCTACGTGAAATCA 960
Qy      961  CAATTTGTGGGAAATGGCGATGCGATGCGAGAGGAATCTTGGCAATTTTGAACGCA 1020
Db      961  CAATTTGTGGGAAATGGCGATGCGATGCGAGAGGAATCTTGGCAATTTTGAACGCA 1020
Qy      1021  TCAACGTTTTCGAGCAATGCGATGATGTAAGACTTAACTTAACTTGTGATTCGCGGT 1080
Db      1021  TCAACGTTTTCGAGCAATGCGATGATGTAAGACTTAACTTAACTTGTGATTCGCGGT 1080
Qy      1081  GCGATTACAGTGGGCGCATCTTCTTGGCGGTCAGCAGCACTTGTGAGGCTTGGCGCT 1140
Db      1081  GCGATTACAGTGGGCGCATCTTCTTGGCGGTCAGCAGCACTTGTGAGGCTTGGCGCT 1140
Qy      1141  GCGACCATGTTTTTGTGGAAGTGTGATCTTTTCTAAACTGCAATAATGCAAAAT 1200
Db      1141  GCGACCATGTTTTTGTGGAAGTGTGATCTTTTCTAAACTGCAATAATGCAAAAT 1200
Qy      1201  CGCGATGCGCTCAATTTGAAGGGGATGGGATTTTGAAGAACTAGAAAGGCTTAG 1260
Db      1201  CGCGATGCGCTCAATTTGAAGGGGATGGGATTTTGAAGAACTAGAAAGGCTTAG 1260
Qy      1261  CAGACAGCGCT 1271
Db      1261  CAGACAGCGCT 1271

```

```

RESULT 3
US-09-471-803A-2
; Sequence 2, Application US/09471803A
; Patent No. 6613545
; GENERAL INFORMATION:
; APPLICANT: KENNERKNECHT, NICOLE
; APPLICANT: SAHM, HERMANN
; APPLICANT: EGGELING, LOTHAR
; APPLICANT: PFEFFERLE, WALTER
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCE CODING FOR THE EXPORT OF
; TITLE OF INVENTION: BRANCHED CHAIN AMINO ACIDS, PROCESS FOR THE
; TITLE OF INVENTION: ISOLATION THEREOF AND USE THEREOF
; FILE REFERENCE: 21123/265486/MAS
; CURRENT APPLICATION NUMBER: US/09/471,803A
; CURRENT FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: DE 199 51 708-8
; PRIOR FILING DATE: 1999-10-27

```

```

; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 2
; LENGTH: 753
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(753)
; OTHER INFORMATION: bmf
; OTHER INFORMATION: ATCC14752
; US-09-471-803A-2

```

```

Query Match      59.2% Score 753; DB 4; Length 753;
Best Local Similarity 100.0%; Pred. No. 6,66-243; Indels 0; Gaps 0;
Matches 753; Conservative 0; Mismatches 0;

```

```

Qy      101  GTGCAAAAACGCAAGAGATTCATTCAAGCTGAGGTGTGCCATCCAGGCGGCTG 160
Db      1  GTGCAAAAACGCAAGAGATTCATTCAAGCTGAGGTGTGCCATCCAGGCGGCTG 60
Qy      161  GAACGAGATGAATGAAGTTATGCGGCTACGAATGCGCAAGTGTAAACCTCCCT 220
Db      61  GAACGAGATGAATGAAGTTATGCGGCTACGAATGCGCAAGTGTAAACCTCCCT 120
Qy      221  GCTGAGGTTTGGGCAATGATACCGATGATGATGATGATGATGATGATGATGATGATG 280
Db      121  GCTGAGGTTTGGGCAATGATACCGATGATGATGATGATGATGATGATGATGATGATG 180
Qy      281  GGTACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 340
Db      181  GGTACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 240
Qy      341  ATGCTGATCATGCGCTCGTGTGATGATGATGATGATGATGATGATGATGATGATGATG 400
Db      241  ATGCTGATCATGCGCTCGTGTGATGATGATGATGATGATGATGATGATGATGATGATG 300
Qy      401  TTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 460
Db      301  TTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 360
Qy      461  AACCCATGCGCGGTTCTATTGATGATGATGATGATGATGATGATGATGATGATGATG 520
Db      361  AACCCATGCGCGGTTCTATTGATGATGATGATGATGATGATGATGATGATGATGATG 420
Qy      521  GCGGCGAGCGCGCGAGCTGTGTGCGTGTGCGATTAATCTCAATGCAATAGCGTTTCA 580
Db      421  GCGGCGAGCGCGCGAGCTGTGTGCGTGTGCGATTAATCTCAATGCAATAGCGTTTCA 480
Qy      581  TCTTACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 640
Db      481  TCTTACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 540
Qy      641  ATTAGGCGCTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 700
Db      541  ATTAGGCGCTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 600
Qy      701  CGAAGCAAAAAGAGATTCCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 760
Db      601  CGAAGCAAAAAGAGATTCCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
Qy      761  GTGTAATTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 820
Db      661  GTGTAATTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 720
Qy      821  ATCCGTAATTCCTTCTTCTGATGATGATGATGATGATGATGATGATGATGATGATG 885
Db      721  ATCCGTAATTCCTTCTTCTGATGATGATGATGATGATGATGATGATGATGATGATG 783

```

```

RESULT 4
US-09-471-803A-4
; Sequence 4, Application US/09471803A

```

```

; Patent No. 6613545
; GENERAL INFORMATION:
; APPLICANT: KENNERKNECHT, NICOLE
; APPLICANT: SAHM, HERMANN
; APPLICANT: EGGELING, LOTAR
; APPLICANT: PEEFERLE, WALTER
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCE CODING FOR THE EXPORT OF
; TITLES OF INVENTION: BRANCHED CHAIN AMINO ACIDS, PROCESS FOR THE
; FILE REFERENCE: 21123/265496/MAS
; CURRENT APPLICATION NUMBER: US/09/471,803A
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: DE 199 51 708-8
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 4
; LENGTH: 324
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(324)
; OTHER INFORMATION: brin
; OTHER INFORMATION: ATCC14752
; US-09-471-803A-4

```

```

Query Match      25.5%; Score 324; DB 4; Length 324;
Best Local Similarity 100.0%; Pred. No. 8.9e-99;
Matches 324; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 853 ATGACACTGATTTCTCCTGATTCCTTTGTTGCGAGATGAGTCACTTACTTT 912
DB 1 ATGACACTGATTTCTCCTGATTCCTTTGTTGCGAGATGAGTCACTTACTTT 60
QY 913 GGGCTCCGGGCGGTCCTTCTTAATCCTTAAGCCCCCTACGATGCAATTTGGGC 972
DB 61 GGGCTCCGGGCGGTCCTTCTTAATCCTTAAGCCCCCTACGATGCAATTTGGGC 120
QY 972 AAAATGGCGATGATGATGCGACAGAGATCTTGCCATTTGACCGCATCAAGTTTGGC 1032
DB 121 AAAATGGCGATGATGATGCGACAGAGATCTTGCCATTTGACCGCATCAAGTTTGGC 180
QY 1033 AGCAATGCGATGATGATGAAAGCTTAACTTTGCTTATGCGCTTGGCATTCAGTG 1092
DB 181 AGCAATGCGATGATGATGAAAGCTTAACTTTGCTTATGCGCTTGGCATTCAGTG 240
QY 1093 GTGGCGCATCTTCTTGGCGATGCGACGACCTTGTGAGCGCTGGCGCTGGCACATCGTT 1152
DB 241 GTGGCGCATCTTCTTGGCGATGCGACGACCTTGTGAGCGCTGGCGCTGGCACATCGTT 300
QY 1153 TTTGTTGAGCTGATGATCTTTTC 1176
DB 301 TTTGTTGAGCTGATGATCTTTTC 324

```

```

RESULT 5
US-09-489-039A-3016
; Sequence 3016, Application US/09489039A
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLES OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; PRIOR FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 3016
; LENGTH: 756
; TYPE: DNA

```

```

; ORGANISM: Klebsiella pneumoniae
; US-09-489-039A-3016

```

```

Query Match      4.6%; Score 58.6; DB 4; Length 756;
Best Local Similarity 52.2%; Pred. No. 3.4e-09;
Matches 130; Conservative 0; Mismatches 119; Indels 0; Gaps 0;

```

```

QY 195 TCGGGAAGGTCATAAAACCTCCCTGTCGACGTTTGGGCAATGACCCGATGTATG 254
DB 62 TCGGGAAGGTCATAAAACCTCCCTGTCGACGTTTGGGCAATGACCCGATGTATG 121
QY 255 CGTTGGTCTCTTGTATTATCAATACGCTACGATGAGGCGACGCCCATGTTTCG 314
DB 122 CTTGCGAGTCATGCTACCCGCGCTGCTTCAACCCGCTGGAAGCCGTCTTCTCT 181
QY 315 GCGTGAATTTGCGGCGCTCCACCGAATGCTGTATGCGCCCTCGTTGGGCGACGCG 374
DB 182 GCATTAATTAAGCGCGCGCGACGAGTTGTATCACCCCATGTTGGCGCGGCGAGTT 241
QY 375 CCTGGGCGCATGCGGCTCACCAATGCTGTGTAATTCGCGCAAGTATTCTATGCGT 434
DB 242 CGCTGCGGTGCGCGCGCTGACGCTGATGCGGATGATGCGGCAATGATGCGGATGCGG 301
QY 435 TTTCATTC 443
DB 302 CTTCACTGC 310

```

```

RESULT 6
US-08-232-463-14
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHRIFFINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-9300
; TELEFAX: (703) 683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear

```

IMMEDIATE SOURCE:
CLONE: PT29C-Fls
US-08-232-463-14

Query Match
Best Local Similarity 3.6%; Pred. No. 28/06;
Matches 14; Conservative 221; Mismatches 157; Indels 0; Gaps 0;

4.1%; Score 52.4; DB 1; Length 7218;
Query Match
Best Local Similarity 3.6%; Pred. No. 28/06;
Matches 14; Conservative 221; Mismatches 157; Indels 0; Gaps 0;

560 TCAATGCAATAGAGTTTCACTCTGATGAGTATCGCGGCTCACCAGAGTGGCATC 619
1050 TCGAGGAGCTTGGGATATATTAAGGAGGCTCGAGGCTTGGCTCTTCTTGTGC 679
620 GCAAGTATGATTCCTTTGAATTAAGGAGGCTCGAGGCTTGGCTCTTCTTGTGC 679
1110 YY 1169
680 ACGGTGACTTGGATTCCTGCGAAGGAGGAGATCCCTTCTGCTGCTGCGAGT 739
1170 YY 1229
740 TTGAGCTTACCATTCCTCTGTGTATTCAGGTGAGGCTTGGCTCTTCTTGTGC 799
1230 YY 1289
800 ATCTCTGAGTCTGTTGACCATCGGATCTTCTTGGGAAAGCTCTAATGACAA 859
1290 YY 1349
860 CTGATTTCTCTGATTCCTCTGTGTGCGAGATGCGAGTCAATTTTGGCTCC 919
1350 YY 1409
920 GGGCGGTCCGTTCTTAATCCTTAAGCCCTA 951
1410 YYYYYYYYYYYYYYYYYYYYYYGTACAA 1441

RESULT 7
US-09-107-532A-1266
Sequence 1266, Application US/09107532A
Patent No. 6583275

GENERAL INFORMATION:

APPLICANT: Lynn A Doucette-Stamm and David Bush

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO

NUMBER OF SEQUENCES: 7310 ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS

CORRESPONDENCE ADDRESS:

ADDRESS: GENE THERAPEUTICS CORPORATION

STREET: 100 Beaver Street

CITY: Waltham

STATE: Massachusetts

COUNTRY: USA

ZIP: 02354

COMPUTER READABLE FORM:

MEDIUM TYPE: CD-ROM ISO9660

COMPUTER: PC

OPERATING SYSTEM: <Unknown>

SOFTWARE: ASCII

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/107,532A

FILING DATE: 30-Jun-1998

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/085,598

FILING DATE: 14 May 1998

APPLICATION NUMBER: 60/051571

FILING DATE: July 2, 1997

ATTORNEY/AGENT INFORMATION:

NAME: Arnieillo, Pamela Deneka

REGISTRATION NUMBER: 40,489

REFERENCE/DOCKET NUMBER: GTC-012

TELECOMMUNICATION INFORMATION:

TELEPHONE: (781)893-5007

TELEFAX: (781)893-8277

INFORMATION FOR SEQ ID NO: 1266:

SEQUENCE CHARACTERISTICS:

LENGTH: 768 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: circular

MOLECULE TYPE: DNA (genomic)

HYPOTHETICAL: NO

ANTI-SENSE: NO

ORIGINAL SOURCE:

ORGANISM: Enterococcus faecium

FEATURE:

NAME/KEY: misc feature

LOCATION: (8) LOCATION 1...768

SEQUENCE DESCRIPTION: SEQ ID NO: 1266:

US-09-107-532A-1266

Query Match
Best Local Similarity 3.1%; Score 39.8; DB 4; Length 768;
Matches 104; Conservative 0; Mismatches 107; Indels 0; Gaps 0;

207 TAAACCTCCCTGCTGAGGTTGGGATGATCCGATGATGATGATGATGATGATGAT 266
92 TCAAGACAGTACTACCTGATTTGGATTAATGATGATGATGATGATGATGATGAT 151
267 TGGTATCAATACGAGTACGATGATGATGATGATGATGATGATGATGATGATGAT 326
152 TTGGAAAGTGGCGGATTTACATCCACTAGTGTGATGATGATGATGATGATGATGAT 211
327 GGGGCTCCAGGAAAGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 386
212 CTGATTCGCGCAATTAATACAGTACGATGATGATGATGATGATGATGATGATGAT 271
387 TCGGCTACACCATTCGATGATGATGATGATGATGATGATGATGATGATGATGAT 417
272 TCGTTTCTGACCTTTCTAGTCAATTCGCG 302

RESULT 8
US-09-252-991A-11376
Sequence 11376, Application US/09252991A
Patent No. 6551795

GENERAL INFORMATION:

APPLICANT: Marc J. Rubenfield et al.

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

NUMBER OF SEQUENCES: 107196,136 AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

FILE REFERENCE: 107196,136

CURRENT APPLICATION NUMBER: US/09/252,991A

PRIOR FILING DATE: 1999-02-18

PRIOR APPLICATION NUMBER: US 60/074,788

PRIOR FILING DATE: 1998-02-18

PRIOR APPLICATION NUMBER: US 60/094,190

PRIOR FILING DATE: 1998-07-27

NUMBER OF SEQ ID NOS: 33142

SEQ ID NO 11376

LENGTH: 1038

TYPE: DNA

ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-11376

Query Match
Best Local Similarity 3.1%; Score 39.2; DB 4; Length 1038;
Matches 77; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

309 TTTCGGCTGATTTTCGGGCTCCACGAATGCTGATGATGATGATGATGATGATGAT 368
473 TGTCCGCAATGCTTTCGGGCTCCACGAATGCTGATGATGATGATGATGATGATGAT 532
369 CAGGCGCTGAGGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 428
533 GTGCGGATCTTTTGTGATCATCTTACACACCTCCCTGACCTGACACACCTGCTCT 592
429 ATGCGTTTTCATTCCTGCTG 448

Db 593 ACGGATGAGCTTCGGGCG 612

RESULT 9

US-09-252-991A-11418
Sequence 11418, Application US/09252991A
Patent No. 6551795

GENERAL INFORMATION:

APPLICANT: Marc J. Rubenfield et al.

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

FILE REFERENCE: 107196.136

CURRENT FILING DATE: 1999-02-18

PRIOR FILING DATE: 1998-02-18

PRIOR APPLICATION NUMBER: US 60/094,190

PRIOR FILING DATE: 1998-07-27

NUMBER OF SEQ ID NOS: 33142

SEQ ID NO 11418

LENGTH: 1086

TYPE: DNA

ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-11418

Query Match 3.1%; Score 39.2; DB 4; Length 1086;

Best Local Similarity 55.0%; Pred. No. 0.015; Mismatches 63; Indels 0; Gaps 0;

Matches 77; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

QY 309 TTTCGGGCTGATTTTCGGGCTCCACGGAATGCTGTCATCGCCCTCGTTGGGCG 368

DB 222 TCTCCGCTCTGCTCTTCCGCGCGCCGCGCCGCTGCTGCGCATGCGCATGCGCAAGGCG 281

QY 369 CAGCGCCCTGCGCGCATGCGCTGCGCATGCGCATGCGCATGCGCATGCGCATGCGCAT 428

DB 282 GTGCGCGGCTTCTTTGATGATCTTCAACACCCCTCTCTGACCTGCGACGCTGCTCT 341

QY 429 ATGCGCTTTCATTTCCGCGCTG 448

DB 342 ACGGATGAGCTTCGGGCG 361

RESULT 10

US-09-621-976-8976/C

Sequence 8976, Application US/09621976

Patent No. 6639063

GENERAL INFORMATION:

APPLICANT: Dumas Milne Edwards, J.B.

APPLICANT: Jobert, S.

APPLICANT: Giordano, J.Y.

TITLE OF INVENTION: ESTs and Encoded Human Proteins.

FILE REFERENCE: GENSET 054PR2

CURRENT FILING DATE: 2000-07-21

CURRENT FILING DATE: 2000-07-21

NUMBER OF SEQ ID NOS: 19335

SEQ ID NO 8976

LENGTH: 399

TYPE: DNA

ORGANISM: Homo sapiens

US-09-621-976-8976

Query Match 3.1%; Score 39; DB 4; Length 399;

Best Local Similarity 12.6%; Pred. No. 0.0086;

Matches 30; Conservative 112; Mismatches 97; Indels 0; Gaps 0;

QY 609 GAGGCGGATGCGAGATGATTCCTTTGAATTAAGGCGCTGAGTTGCCCTTGGT 668

DB 242 SAAMRKAKMTGWSRSGSRSTGYAMTKSKCTSRKMYKKRRKKRCKTSTXT 163

QY 669 CTCTCTTGTACCGCTGACTTGTGATTCCTGCGAAGAAAAGCAGATCCCTTCTGCG 728

DB 192 CTGSGTYCKAKAYTKKRRKTRMTYYKYKSMKTKMRKATAYTTRKMTETKWT 123

QY 729 TGCTGCGAGTTTGTAGCTTACCATGCTCTTGATTAATCCAGTGAAGCCCTATTGG 788

DB 122 CTGCKCTTGTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTG 63

QY 789 CCGCGCTGCTGATCTTCTTGGGCTGCTGATGACATCGGATCTTCTTGGGAAAGCT 847

DB 62 MMKMGKMMKMTYYTYTMMKMSKTYTMMSCYARKCRYYATYTTMTGCTGT 4

RESULT 11

US-09-621-976-18033

Sequence 18033, Application US/09621976

Patent No. 6639063

GENERAL INFORMATION:

APPLICANT: Dumas Milne Edwards, J.B.

APPLICANT: Jobert, S.

APPLICANT: Giordano, J.Y.

TITLE OF INVENTION: ESTs and Encoded Human Proteins.

FILE REFERENCE: GENSET 054PR2

CURRENT FILING DATE: 2000-07-21

NUMBER OF SEQ ID NOS: 19335

SOFTWARE: Patent.pm

SEQ ID NO 18033

LENGTH: 474

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: misc_feature

LOCATION: 16

OTHER INFORMATION: n=a, g, c or t

US-09-621-976-18033

Query Match 2.9%; Score 37; DB 4; Length 474;

Best Local Similarity 14.4%; Pred. No. 0.045; Mismatches 156; Indels 0; Gaps 0;

Matches 49; Conservative 135; Mismatches 156; Indels 0; Gaps 0;

QY 522 CGGCGAGCCCGAGGCTGCTGCGCGGACTTATCTCATGCAATGCAATGCGTTCACT 561

DB 13 YSGNCTCYCMARKMSYKXGSMYKSSCYKSCMMCMKSYKSGSYKTTTWTAW 72

QY 582 CTTACTGGTATTTGCGGCTCTCACCGAGATGCGGAGTGCAGATGATTCCTTTGAA 641

DB 73 WMTTTRGKXARRMSGSGKTYTMMCSKKTYSCKAGRWKGYYSWYTCYKACTY 132

QY 642 TTAAGGCTGCGAGTTGCGCTTGTCTCTCTTGTCACTGACTTGTGATTCCTGCC 701

DB 133 WKRNYCSCCMYTKGSGSWTTTMMRRKSKYKRTKGGKKTMMMAACYTWR 192

QY 702 GAACGAAAAGACAGATCCCTTCTGCTGCTGCTGCGAGTTGAGCTTACATGCTCTG 761

DB 193 YMMMMRRAAAATKYCCMMSKTMCWACCMMCCMRARSCMRSCMRSYTMMCTYY 252

QY 762 TGTAAATTCAGATCAGGCGCTATTTGCGGCGCTGATCTTCTTGGCTGTGACCA 821

DB 253 MMTGGRMYTMMRGMMKMTMYKSKSMKSGCMRBAWARTTYTAMYYTYTKR 312

QY 822 TCCGTAATCTTCTTGGGAAAAGCTGCTAATGACACT 861

DB 313 MCTYMRKTYCMMMYSRWRSMTARFAMMMCMYMY 352

RESULT 12

US-09-252-991A-3205/C

Sequence 3205, Application US/09252991A

Patent No. 6551795

GENERAL INFORMATION:

APPLICANT: Marc J. Rubenfield et al.

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

FILE REFERENCE: 107196.136

CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 3205
LENGTH: 1140
TYPE: DNA
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-3205

Query Match 2.9%; Score 36.8; DB 4; Length 1140;
Best Local Similarity 58.0%; Pred. No. 0.097;
Matches 65; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

QY 311 TCCGCGCTGATTTCGCGCGGCTCCACGGAATGCTGTCATCGCCCTGTTGGGCGCA 370
DB 465 TACTTCCGATCAACCTCGCGCGCCCTCGCGCGCTGATCGGCTGACTCTCGCGCTC 406
QY 371 GCGCGCCGCGCGCGCATCGCGCTCACCATTCGTCGTAAGTCCGCGCAG 422
DB 405 GCGCGCGCGCGCGCATCTTCTGTCACCGCGCTGTCCTTCTGTCAG 354

RESULT 13
US-09-252-991A-2761
Sequence 2761, Application US/09252991A
Patent No. 6551795

GENERAL INFORMATION:
APPLICANT: Marc J. Rudenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 2761
LENGTH: 2061
TYPE: DNA
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-2761

Query Match 2.9%; Score 36.8; DB 4; Length 2061;
Best Local Similarity 58.0%; Pred. No. 0.15;
Matches 65; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

QY 311 TCCGCGCTGATTTCGCGCGGCTCCACGGAATGCTGTCATCGCCCTGTTGGGCGCA 370
DB 1273 TACTTCCGATCAACCTCGCGCGCCCTCGCGCGCTGATCGGCTGACTCTCGCGCTC 1332
QY 371 GCGCGCCGCGCGCGCATCGCGCTCACCATTCGTCGTAAGTCCGCGCAG 422
DB 1333 GCGCGCGCGCGCGCATCTTCTGTCACCGCGCTGTCCTTCTGTCAG 1384

RESULT 14
US-09-020-956-12/c
Sequence 12, Application US/09020956
Patent No. 6261562

GENERAL INFORMATION:
APPLICANT: Xu, Jiangchun
TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY OF PROSTATE CANCER AND METHODS FO
NUMBER OF SEQUENCES: 178
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue

CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98104

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/020,956
FILING DATE: 09-FEB-1998
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Maki, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 210121.427C2

TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 751 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-09-020-956-12

Query Match 2.9%; Score 36.4; DB 3; Length 751;
Best Local Similarity 59.8%; Pred. No. 0.099;
Matches 58; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

QY 327 CCGGCTCCACGGAATGCTGTCATCGCCCTGTTGGGCGCGCAGCCCTGGGCGCA 386
DB 410 CGTCTTCTTCATCNCCTCCTCATCTTCATTCGTCGAGGTCAGCTGTCGTCGCT 351
QY 387 TCCGCGTCCACCATTCGTCGTAAGTCCGCGCAGT 423
DB 350 TGGTGTACACCAATGCTGAGACACTTCCCTGAGCT 314

RESULT 15
US-09-030-607-12/c

Sequence 12, Application US/09030607
Patent No. 6262245
GENERAL INFORMATION:

APPLICANT: Xu, Jiangchun
TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY OF PROSTATE CANCER AND METHODS FO
NUMBER OF SEQUENCES: 224
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98104

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/030,607
FILING DATE: 25-FEB-1998
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Maki, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 210121.427C3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900

TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 751 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-09-030-607-12

Query Match 2.9%; Score 36.4; DB 3; Length 751;
Best Local Similarity 59.8%; Pred. No. 0.099;
Matches 58; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

Qy 327 CGGGCTCCACGGAATGCTGATCGCCCTGTTGTGGGCGCAGGCCCTGGGGCCCA 386
Db 410 CGTTCTTCTTCATCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCT 351
Qy 387 TCGCGCTCCACGGAATGCTGATCGCCCTGTTGTGGGCGCAGGCCCTGGGGCCCA 423
Db 350 TGGTGTACACGGAATGCTGATCGCCCTGTTGTGGGCGCAGGCCCTGGGGCCCA 314

Search completed: April 19, 2004, 18:38:03
Job time: 113.018 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 17, 2004, 06:48:18 ; Search time 617.884 Seconds
(without alignments)
9204.305 Million cell updates/sec

Title: US-10-608-504-1

Perfect score: 1271
Sequence: 1 gcgcgcacacagcagcctcag.....agccttaacagcagcgcct 1271

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 2890132 seqs, 2237290429 residues

Total number of hits satisfying chosen parameters: 5780264

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications_NA:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/1/pubpna/PCF_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/1/pubpna/PCFUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
- 14: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/1/pubpna/US10C_NEW_PUB.seq:*
- 17: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
- 18: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1271	100.0	1271	15	US-10-608-504-1
2	1271	100.0	1271	15	US-10-608-504-1
3	1271	100.0	3309400	9	US-09-738-626-1
4	753	59.2	753	9	US-09-738-626-289
5	753	59.2	753	15	US-10-608-504-2
6	365.6	28.8	397	12	US-10-134-640-7
7	365.6	28.8	778	12	US-10-134-640-9
8	324	25.5	324	9	US-09-738-626-250
9	324	25.5	324	15	US-10-608-504-4
10	185	14.6	715	12	US-10-134-640-1
11	127.4	10.0	711	9	US-09-738-626-3408
12	127.4	10.0	3309400	9	US-09-738-626-1
13	43	3.4	594	14	US-10-156-761-6067
14	43	3.4	9025608	14	US-10-156-761-1
15	41.4	3.3	351	14	US-10-184-644-80

C 16	41.4	3.3	351	14	US-10-184-634-80	Sequence 80, Appl
C 17	40	3.1	671	14	US-10-184-644-346	Sequence 346, App
C 18	40	3.1	671	14	US-10-184-634-346	Sequence 346, App
C 19	39.8	3.1	596	14	US-10-184-644-310	Sequence 310, App
C 20	39.8	3.1	596	14	US-10-184-634-310	Sequence 310, App
C 21	39.8	3.1	596	14	US-10-063-685-100	Sequence 100, App
C 22	38.4	3.0	594	12	US-10-142-426-10	Sequence 10, Appl
C 23	38.4	3.0	594	14	US-10-123-153-10	Sequence 10, Appl
C 24	38.4	3.0	594	14	US-10-146-731-10	Sequence 10, Appl
C 25	38.4	3.0	594	14	US-10-140-472-10	Sequence 10, Appl
C 26	38.4	3.0	594	14	US-10-141-761-10	Sequence 10, Appl
C 27	38.4	3.0	594	14	US-10-142-885-10	Sequence 10, Appl
C 28	38.4	3.0	594	14	US-10-158-790-10	Sequence 10, Appl
C 29	38.4	3.0	594	15	US-10-137-871-10	Sequence 10, Appl
C 30	38.4	3.0	594	15	US-10-140-922-10	Sequence 10, Appl
C 31	38.4	3.0	594	15	US-10-141-736-10	Sequence 10, Appl
C 32	38.4	3.0	594	15	US-10-141-759-10	Sequence 10, Appl
C 33	38.4	3.0	594	15	US-10-141-805-10	Sequence 10, Appl
C 34	38.4	3.0	594	15	US-10-140-864-10	Sequence 10, Appl
C 35	38.2	3.0	946	9	US-09-778-320-211	Sequence 211, App
C 36	38.2	3.0	946	9	US-09-910-689-211	Sequence 211, App
C 37	38.2	3.0	946	13	US-10-010-742-211	Sequence 211, App
C 38	37.8	3.0	457	14	US-10-184-644-48	Sequence 48, Appl
C 39	37.8	3.0	457	14	US-10-184-634-48	Sequence 48, Appl
C 40	37.8	3.0	457	14	US-10-063-685-12	Sequence 12, Appl
C 41	37.4	2.9	557	14	US-10-184-644-326	Sequence 326, App
C 42	37.4	2.9	557	14	US-10-184-634-326	Sequence 326, App
C 43	37	2.9	186	14	US-10-156-761-5719	Sequence 5719, Ap
C 44	37	2.9	9025608	14	US-10-156-761-1	Sequence 1, Appl
C 45	36.8	2.9	1230	12	US-10-282-122A-30075	Sequence 30075, A

ALIGNMENTS

RESULT 1
US-10-608-504-1
Sequence 1, Application US/10608504
Publication No. US20040014123A1
GENERAL INFORMATION:
APPLICANT: KENNERNECHT, NICOLE
APPLICANT: SAHM, HERMANN
APPLICANT: BEGELING, LOTMAR
APPLICANT: PFEFFERLE, WALTER
TITLE OF INVENTION: NUCLEOTIDE SEQUENCE CODING FOR THE EXPORT OF
TITLE OF INVENTION: BRANCHED CHAIN AMINO ACIDS, PROCESS FOR THE
FILE REFERENCE: 7601/80525
CURRENT APPLICATION NUMBER: US/10/608,504
CURRENT FILING DATE: 2003-06-30
PRIOR APPLICATION NUMBER: US 09/471,803
PRIOR FILING DATE: 1999-12-23
PRIOR APPLICATION NUMBER: DE 199 51 708.8
PRIOR FILING DATE: 1999-10-27
NUMBER OF SEQ ID NOS: 12
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1
LENGTH: 1271
TYPE: DNA
ORGANISM: Corynebacterium glutamicum
FEATURE:
NAME/KEY: gene
LOCATION: (101)..(853)
OTHER INFORMATION: brnf
FEATURE:
NAME/KEY: gene
LOCATION: (853)..(1176)
OTHER INFORMATION: brns
FEATURE:
OTHER INFORMATION: ATCC14752
US-10-608-504-1
Query Match 100.0%, Score 1271, DB 15, Length 1271;

Best Local Similarity 100.0%; Pred. No. 0; Matches 1271; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 GGGGATCAATGGAATCTGATTCATATATGTCACAAATAGCCTAGTGAAGTGGCGCAAC 60
Db 1 GGGGATCAATGGAATCTGATTCATATATGTCACAAATAGCCTAGTGAAGTGGCGCAAC 60
QY 61 TGGCAACAAACTACCCGGCAATGNGATATGATGTGTGCAAAAACGCAAGAT 120
Db 61 TGGCAACAAACTACCCGGCAATGNGATATGATGTGTGCAAAAACGCAAGAT 120
QY 121 TCATTCAAGCCTGAGAGTGTGCGCATCCAGGAGCCTGGAACCAATGATTAAGTTA 180
Db 121 TCATTCAAGCCTGAGAGTGTGCGCATCCAGGAGCCTGGAACCAATGATTAAGTTA 180
QY 181 TGGGCGCTACGAATATGCGGCAAGGTCTAAAACTCCCTGTGTGAGGTTTGGGATGTA 240
Db 181 TGGGCGCTACGAATATGCGGCAAGGTCTAAAACTCCCTGTGTGAGGTTTGGGATGTA 240
QY 241 CCGATTTGATATGCGTTGATCTCTTGTATTCATATGAGGCTACGAATGATGGGAGC 300
Db 241 CCGATTTGATATGCGTTGATCTCTTGTATTCATATGAGGCTACGAATGATGGGAGC 300
QY 301 CCGACTGTTTTCCGCTGATTTTGGGCGCTCCACCGAATGCTGTGATGCCCCCTGT 360
Db 301 CCGACTGTTTTCCGCTGATTTTGGGCGCTCCACCGAATGCTGTGATGCCCCCTGT 360
QY 361 TGTGGGCGAGGCGCCCTGAGGCGCATGCGGCTCACCATTTGTTGAACTTCCGCA 420
Db 361 TGTGGGCGAGGCGCCCTGAGGCGCATGCGGCTCACCATTTGTTGAACTTCCGCA 420
QY 421 CGTATTCATAGCGTTTCAATCCCGTGCATGTGTCAAAAACCCCATTTCCCGTTCTA 480
Db 421 CGTATTCATAGCGTTTCAATCCCGTGCATGTGTCAAAAACCCCATTTCCCGTTCTA 480
QY 481 TTGCGTTTCCGCTATAGAGAGCCATCCAGCATGTGGGCGAGGCGCGAGCTG 540
Db 481 TTGCGTTTCCGCTATAGAGAGCCATCCAGCATGTGGGCGAGGCGCGAGCTG 540
QY 541 GTGCGGCTGCGCATTAATCTCAATGCAATAGCGTTTCACTCTAATGCGGATTCGCGG 600
Db 541 GTGCGGCTGCGCATTAATCTCAATGCAATAGCGTTTCACTCTAATGCGGATTCGCGG 600
QY 601 TCTCACCGGAGTGGCGATGCGAGATGTATCTCTTTGAAATTAAGGCGCTGAGTGGC 660
Db 601 TCTCACCGGAGTGGCGATGCGAGATGTATCTCTTTGAAATTAAGGCGCTGAGTGGC 660
QY 661 CCTTGTCTCTCTTTGTCAAGCTGATTTGATTCCTGCGAAGCAAAACAGATCCC 720
Db 661 CCTTGTCTCTCTTTGTCAAGCTGATTTGATTCCTGCGAAGCAAAACAGATCCC 720
QY 721 TTCTGTGCTGTGCGAGGTTTGAAGCTTACCATTTGCTTTGTGATTCAGGTCAGGC 780
Db 721 TTCTGTGCTGTGCGAGGTTTGAAGCTTACCATTTGCTTTGTGATTCAGGTCAGGC 780
QY 781 CCTATTGCGGCGCTGATCTCTTGGATGCTGTGACCATCCGTAATCTTCTTGGG 840
Db 781 CCTATTGCGGCGCTGATCTCTTGGATGCTGTGACCATCCGTAATCTTCTTGGG 840
QY 841 AAAAGCTGCTAAATGACATGATTTCTCTGTATTTCTCTTGTGTCAGATATGTA 900
Db 841 AAAAGCTGCTAAATGACATGATTTCTCTGTATTTCTCTTGTGTCAGATATGTA 900
QY 901 GTCAATTAATTTTGGCGCTCCGGGCGGTTCCGTTCTTAATCTTAAGCCCTAGCTGATCA 960
Db 901 GTCAATTAATTTTGGCGCTCCGGGCGGTTCCGTTCTTAATCTTAAGCCCTAGCTGATCA 960
QY 961 CAATTTGTGGGCAAAATGGCGATGTGATGCGAGCAAGATCTTTGACCGCA 1020
Db 961 CAATTTGTGGGCAAAATGGCGATGTGATGCGAGCAAGATCTTTGACCGCA 1020
QY 1021 TCAAGCTTTGGGCAATGGGATGATGATGTAAGATCTTAACCTTTGTGTCATGCGCTT 1080

```

Query Match 100.0%; Score 1271; DB 15; Length 1271; Best Local Similarity 100.0%; Pred. No. 0; Matches 1271; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Db 1021 TCAAGCTTTGGGCAATGGGATGATGATGTAAGATCTTAACCTTTGTGTCATGCGCTT 1080
QY 1081 GCGATTACAGTGTGGCGCATCTTTTGGGCGTGCAGCACTTTGATGAGCGTGGCGCT 1140
Db 1081 GCGATTACAGTGTGGCGCATCTTTTGGGCGTGCAGCACTTTGATGAGCGTGGCGCT 1140
QY 1141 GGCACCATGCTTTTGTGATGATGTAATCTTTCTAAACCTGATTAATACAAAT 1200
Db 1141 GGCACCATGCTTTTGTGATGATGTAATCTTTCTAAACCTGATTAATACAAAT 1200
QY 1201 CCGATGCGCTCAATTTGAAAGGAGATGCGGATTTTAAAGAACTTGAAGGCTTAAG 1260
Db 1201 CCGATGCGCTCAATTTGAAAGGAGATGCGGATTTTAAAGAACTTGAAGGCTTAAG 1260
QY 1261 CAGACAGCGCT 1271
Db 1261 CAGACAGCGCT 1271

RESULT 2
US-10-608-504-6
; Sequence 6, Application US/10608504
; Publication No. US20040014123A1
; GENERAL INFORMATION:
; APPLICANT: KENNERNBECHT, NICOLE
; APPLICANT: SAMM, HERMANN
; APPLICANT: EGGELE, LOTMAR
; APPLICANT: PFEFFERLE, WALTER
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCE CODING FOR THE EXPORT OF
; TITLE OF INVENTION: BRANCHED CHAIN AMINO ACIDS, PROCESS FOR THE
; TITLE OF INVENTION: ISOLATION THEREOF AND USE THEREOF
; FILE REFERENCE: 7601/80525
; CURRENT APPLICATION NUMBER: US/10/608,504
; PRIOR FILING DATE: 2003-06-30
; PRIOR APPLICATION NUMBER: US 09/471,803
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: DE 199 51 708.8
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 1271
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
; FEATURES:
; NAME/KEY: gene
; LOCATION: (101)..(853)
; OTHER INFORMATION: brnF
; FEATURE:
; NAME/KEY: gene
; LOCATION: (853)..(1176)
; OTHER INFORMATION: brnE
; FEATURE:
; OTHER INFORMATION: ATCC13032
US-10-608-504-6

Query Match 100.0%; Score 1271; DB 15; Length 1271; Best Local Similarity 100.0%; Pred. No. 0; Matches 1271; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Db 181 TCGGCGCTACGAAATCGCGAAGCTTAAACCTCCCTTCTCAGGTTGGGCAATGA 240
 Qy 241 CCCGATTGGTATGTTGGTCTCTTGGTATTCATATACGATAGGATGGTGGGAGC 300
 Db 241 CCCGATTGGTATGTTGGTCTCTTGGTATTCATATACGATAGGATGGTGGGAGC 300
 Qy 301 CCCGATTGGTATGTTGGTCTCTTGGTATTCATATACGATAGGATGGTGGGAGC 360
 Db 301 CCCGATTGGTATGTTGGTCTCTTGGTATTCATATACGATAGGATGGTGGGAGC 360
 Qy 361 TGTGGGCGAGGCGCCCTGGGCGCATGCGCTACCACTTGGTGGTGAATCCGCGCA 420
 Db 361 TGTGGGCGAGGCGCCCTGGGCGCATGCGCTACCACTTGGTGGTGAATCCGCGCA 420
 Qy 421 CGTATTCATAGCGTTTCAATCCCGCTCATGTGGTCAAAAACCCCATTTGCCGTTCTA 480
 Db 421 CGTATTCATAGCGTTTCAATCCCGCTCATGTGGTCAAAAACCCCATTTGCCGTTCTA 480
 Qy 481 TTGGGTTTTCGGCTTATCGAGAAAGCTACGACATCTCGGCGCAAGCCCGGAGGCTG 540
 Db 481 TTGGGTTTTCGGCTTATCGAGAAAGCTACGACATCTCGGCGCAAGCCCGGAGGCTG 540
 Qy 541 GTCCGCGTGGGAGCTTATCTCATATGCAATAGCGTTTCACTCTACTGGTATTCGGCGG 600
 Db 541 GTCCGCGTGGGAGCTTATCTCATATGCAATAGCGTTTCACTCTACTGGTATTCGGCGG 600
 Qy 601 TCTCACCGGAGTGGGAGTGGGAGTGGTATTCCTTTGAATTAAGGGCTTCGAGTTGCG 660
 Db 601 TCTCACCGGAGTGGGAGTGGGAGTGGTATTCCTTTGAATTAAGGGCTTCGAGTTGCG 660
 Qy 661 CTTTGGCTCTCTCTTGTCAAGCTGATTCGATTCCTGCGGAAAGCAAAAGCAGATGCC 720
 Db 661 CTTTGGCTCTCTCTTGTCAAGCTGATTCGATTCCTGCGGAAAGCAAAAGCAGATGCC 720
 Qy 721 TTCTCTGCTCTCGAGGTTTGAAGCTTCAACATGCTCTTGTGTATTCAGGTTCAAGG 780
 Db 721 TTCTCTGCTCTCGAGGTTTGAAGCTTCAACATGCTCTTGTGTATTCAGGTTCAAGG 780
 Qy 781 CCTATTGCGGCGCTGATCTTCTTGGGTCGTGTCACACCGGATCTCTTGGG 840
 Db 781 CCTATTGCGGCGCTGATCTTCTTGGGTCGTGTCACACCGGATCTCTTGGG 840
 Qy 841 AAAGGCTCTAAATGACAACTGATTTCTCTGATTTCTCTTGTGTGCAATATGCA 900
 Db 841 AAAGGCTCTAAATGACAACTGATTTCTCTGATTTCTCTTGTGTGCAATATGCA 900
 Qy 901 GTCAATTCTTTTGGCGCTCGGCGCGCTTCCGTTTAACTTAAAGCCCTTACGTAATCA 960
 Db 901 GTCAATTCTTTTGGCGCTCGGCGCGCTTCCGTTTAACTTAAAGCCCTTACGTAATCA 960
 Qy 961 CAATTGTTGGGCAAAATGCGATGTGATGTCAGAGAAATCTTGGCAATTTGACGCA 1020
 Db 961 CAATTGTTGGGCAAAATGCGATGTGATGTCAGAGAAATCTTGGCAATTTGACGCA 1020
 Qy 1021 TCAAGGTTTTCGAGCAATGCGATGATGTAAGATCTTAACTTTGGTCTATTCGCGTT 1080
 Db 1021 TCAAGGTTTTCGAGCAATGCGATGATGTAAGATCTTAACTTTGGTCTATTCGCGTT 1080
 Qy 1081 GCGATTACAGTGTGGCGATCTTCTTGGCGGTGACGACACTTGTGAAGGCTGGCGCT 1140
 Db 1081 GCGATTACAGTGTGGCGATCTTCTTGGCGGTGACGACACTTGTGAAGGCTGGCGCT 1140
 Qy 1141 GCGACCATCGTTTGTGGTGAATCTTTTCTAAACCTGATTAATAACAAAT 1200
 Db 1141 GCGACCATCGTTTGTGGTGAATCTTTTCTAAACCTGATTAATAACAAAT 1200
 Qy 1201 CCGGATGCGCTCAATTTGAAGGCGATTTGTAAGGAACCTTAAGAAAGGCTTAAG 1260
 Db 1201 CCGGATGCGCTCAATTTGAAGGCGATTTGTAAGGAACCTTAAGAAAGGCTTAAG 1260
 Qy 1261 CAGACAGCGCT 1271
 Db 1261 CAGACAGCGCT 1271

Db 1261 CAGACAGCGCT 1271
 RESULT 3
 US-09-738-626-1
 ; Sequence 1, Application US/09738626
 ; Publication No. US20020197605A1
 ; GENERAL INFORMATION:
 ; APPLICANT: NAKAGAWA, SATOSHI
 ; APPLICANT: MIZOGUCHI, HIROSHI
 ; APPLICANT: ANDO, SEIKO
 ; APPLICANT: HAYASHI, MIKIRO
 ; APPLICANT: OCHIAI, KEIKO
 ; APPLICANT: YOKOI, HARUHIKO
 ; APPLICANT: TATEISHI, NAOKO
 ; APPLICANT: SENOH, AKIHIRO
 ; APPLICANT: IKEDA, MASATO
 ; APPLICANT: OZAKI, AKIO
 ; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
 ; FILE REFERENCE: 249-125
 ; CURRENT APPLICATION NUMBER: US/09/738, 626
 ; CURRENT FILING DATE: 2000-12-18
 ; PRIOR APPLICATION NUMBER: JP 99/377484
 ; PRIOR FILING DATE: 1999-12-16
 ; PRIOR APPLICATION NUMBER: JP 00/159162
 ; PRIOR FILING DATE: 2000-04-07
 ; PRIOR APPLICATION NUMBER: JP 00/280988
 ; PRIOR FILING DATE: 2000-08-03
 ; NUMBER OF SEQ ID NOS: 7059
 ; SOFTWARE: PatentIn ver. 3.0
 ; SEQ ID NO 1
 ; LENGTH: 3309400
 ; TYPE: DNA
 ; ORGANISM: Corynebacterium glutamicum
 ; US-09-738-626-1
 Query Match 100.0%; Score 1271; DB 9; Length 3309400;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1271; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 GCGGATTCATAGCAATCTAGCTTCATATATTCGACAAATAGCTAGTGGGCGCAAC 60
 Db 276729 GCGGATTCATAGCAATCTAGCTTCATATATTCGACAAATAGCTAGTGGGCGCAAC 276788
 Qy 61 TGGCAACAAACTACCGGCAATTTGTGATGATTTAGTGTGCAAAAAACGCAAGAT 120
 Db 276789 TGGCAACAAACTACCGGCAATTTGTGATGATTTAGTGTGCAAAAAACGCAAGAT 276848
 Qy 121 TCATTCAAGCTTGAAGTGTGCGCATCCCAAGGAGCCCTGGAAACGATGATTAAGTTA 180
 Db 276849 TCATTCAAGCTTGAAGTGTGCGCATCCCAAGGAGCCCTGGAAACGATGATTAAGTTA 276908
 Qy 181 TCGGCGCTACGAATGCGCGCAAGGCTTAAACCTCTCTGTCAGGTTTGGGCAATGA 240
 Db 276909 TCGGCGCTACGAATGCGCGCAAGGCTTAAACCTCTCTGTCAGGTTTGGGCAATGA 276968
 Qy 241 CCCGATTGGTATGTTGGTCTCTTGGTATTCATATACGATAGGATGGTGGGAGC 300
 Db 276969 CCCGATTGGTATGTTGGTCTCTTGGTATTCATATACGATAGGATGGTGGGAGC 277028
 Qy 301 CCCGATTGGTATGTTGGTCTCTTGGTATTCATATACGATAGGATGGTGGGAGC 360
 Db 277029 CCCGATTGGTATGTTGGTCTCTTGGTATTCATATACGATAGGATGGTGGGAGC 277088
 Qy 361 TGTGGGCGAGCGCCCTGGGCGCATGCGCTACCACTTGGTGGTGAATCCGCGCA 420
 Db 277089 TGTGGGCGAGCGCCCTGGGCGCATGCGCTACCACTTGGTGGTGAATCCGCGCA 277148
 Qy 421 CGTATTCATAGCGTTTCAATCCCGCTCATGTGGTCAAAAACCCCATTTGCCGTTCTA 480
 Db 277149 CGTATTCATAGCGTTTCAATCCCGCTCATGTGGTCAAAAACCCCATTTGCCGTTCTA 277208
 Qy 481 TTGGGTTTTCGGCTTATCGAGAAAGCTACGACATCTCGGCGCAAGCCCGGAGGCTG 540

Db 277209 TTGGTTTGGGCTTATGACGAAAGCTTACGAGCTCACTCGGCGAGCCCGAGGCTG 277268
Qy 541 GTCCGCGTGGGACTTATCTCAATGCAATAGCGTTTCACTCTACTGAGTATTCGCGG 600
Db 277269 GTCCGCGTGGGACTTATCTCAATGCAATAGCGTTTCACTCTACTGAGTATTCGCGG 277328
Qy 601 TCTCACCGGAGTGGGAGTGGGAGTATTCCTTTTAAATTTAAGGCGCTCGAGTTCCG 660
Db 277329 TCTCACCGGAGTGGGAGTGGGAGTATTCCTTTTAAATTTAAGGCGCTCGAGTTCCG 277388
Qy 661 CTTTGGCTCTCTCTCTGTCAAGCTGACTTGTGATTCCTCGCGAAAGAAAGAGATCCC 720
Db 277389 CTTTGGCTCTCTCTCTGTCAAGCTGACTTGTGATTCCTCGCGAAAGAAAGAGATCCC 277448
Qy 721 TTTCTGCTGCTGCGAGGTTTGAAGTTTCAACATTCCTTGTGTATTCAGAGTCAAGC 780
Db 277449 TTTCTGCTGCTGCGAGGTTTGAAGTTTCAACATTCCTTGTGTATTCAGAGTCAAGC 277508
Qy 781 CCTATTTGGGCGGCTGCTGATCTTCTGAGTCTGTGACCAATCCGAGTACTTCTTGGG 840
Db 277509 CCTATTTGGGCGGCTGCTGATCTTCTGAGTCTGTGACCAATCCGAGTACTTCTTGGG 277568
Qy 841 AAAGCTGCTAAATGACAACTGATTTCTCTGATTTCTCTGTTTGTGCGAGATGCGA 900
Db 277569 AAAGCTGCTAAATGACAACTGATTTCTCTGATTTCTCTGTTTGTGCGAGATGCGA 277628
Qy 901 GTCAATTGTTGGGCTCGGCGGCTCGGTTCTTAATCCTTAAGCCCTTAAGGATCA 960
Db 277629 GTCAATTGTTGGGCTCGGCGGCTCGGTTCTTAATCCTTAAGCCCTTAAGGATCA 277688
Qy 961 CAATTGTTGGGCAAAATGGCGATGTGATGCCAGAGGAAATCCTTGCAATTTGACGCA 1020
Db 277689 CAATTGTTGGGCAAAATGGCGATGTGATGCCAGAGGAAATCCTTGCAATTTGACGCA 277748
Qy 1021 TTAAGTTTGGGCAAAATGGCGATGTGATGCCAGAGGAAATCCTTGCAATTTGACGCT 1080
Db 277749 TTAAGTTTGGGCAAAATGGCGATGTGATGCCAGAGGAAATCCTTGCAATTTGACGCT 277808
Qy 1081 GCGATTACAGTGTGGCGCATCTTCTGCGGCTGACGACGACCTTGTGAGGCTTGGCCT 1140
Db 277809 GCGATTACAGTGTGGCGCATCTTCTGCGGCTGACGACGACCTTGTGAGGCTTGGCCT 277868
Qy 1141 GCGATTACAGTGTGGCGCATCTTCTGCGGCTGACGACGACCTTGTGAGGCTTGGCCT 1200
Db 277869 GCGATTACAGTGTGGCGCATCTTCTGCGGCTGACGACGACCTTGTGAGGCTTGGCCT 277928
Qy 1201 CCGCATGCGCTCAATTTGAAGGAGATGCGGATTTTAAAGAACTTAAGAAAGGCTTAAG 1260
Db 277929 CCGCATGCGCTCAATTTGAAGGAGATGCGGATTTTAAAGAACTTAAGAAAGGCTTAAG 277988
Qy 1261 CAGACAGCGCT 1271
Db 277989 CAGACAGCGCT 277999

RESULT 4
US-09-738-626-289
Sequence 289, Application US/09738626
GENERAL INFORMATION:
APPLICANT: NAKAGAWA, SATOSHI
APPLICANT: MIZOGUCHI, HIROSHI
APPLICANT: ANDO, SEIKO
APPLICANT: HAYASHI, MIKIRO
APPLICANT: OCHIAI, KEIKO
APPLICANT: YOKOI, HARUHIKO
APPLICANT: TATEISHI, NAKO
APPLICANT: SENOH, AKIHIRO
APPLICANT: IKEDA, MASATO
APPLICANT: OZAKI, AKIO
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-125

CURRENT APPLICATION NUMBER: US/09/738, 626
CURRENT FILING DATE: 2000-12-18
PRIOR APPLICATION NUMBER: JP 99/377484
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: JP 00/159162
PRIOR FILING DATE: 2000-04-07
PRIOR APPLICATION NUMBER: JP 00/280988
PRIOR FILING DATE: 2000-08-03
NUMBER OF SEQ ID NOS: 7059
SOFTWARE: PatentIn ver. 3.0
SEQ ID NO: 289
LENGTH: 753
TYPE: DNA
ORGANISM: Corynebacterium glutamicum
US-09-738-626-289

Query Match 59.2%; Score 753; DB 9; Length 753;
Best Local Similarity 100.0%; Pred. No. 6,5e-245; Indels 0; Gaps 0;
Matches 753; Conservative 0; Mismatches 0;

Qy 101 GTGCAAAAAACGACAGAGATTCAATCAAGCTGAGAGTGTGCGCATCCAAAGCAGCCCTG 160
Db 1 GTGCAAAAAACGACAGAGATTCAATCAAGCTGAGAGTGTGCGCATCCAAAGCAGCCCTG 60
Qy 161 GAAACGATGATTAAGGTTATGCGGCTACGAAATGCGCAAGTTTAAAACTTCCCTT 220
Db 61 GAAACGATGATTAAGGTTATGCGGCTACGAAATGCGCAAGTTTAAAACTTCCCTT 120
Qy 221 GCTGAGGTTTGGGCGATGACCGGATTTGATTTGATTTGATTTGATTTGATTTGATTTG 280
Db 121 GCTGAGGTTTGGGCGATGACCGGATTTGATTTGATTTGATTTGATTTGATTTGATTTG 180
Qy 281 GAGTACGATGATGAGGAGGCGGCGATTTTCCGCGCTGATTTTCCGCGGCTTCCACGAA 340
Db 181 GAGTACGATGATGAGGAGGCGGCGATTTTCCGCGCTGATTTTCCGCGGCTTCCACGAA 240
Qy 341 ATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 400
Db 241 ATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300
Qy 401 TTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 460
Db 301 TTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 360
Qy 461 AAGCCCATGCGGCTTCTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 520
Db 361 AAGCCCATGCGGCTTCTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420
Qy 521 GCGGCGAGGCGCGAGGCTGTGCGGCGAGGCTTATGATGATGATGATGATGATGATGATGAT 580
Db 421 GCGGCGAGGCGCGAGGCTGTGCGGCGAGGCTTATGATGATGATGATGATGATGATGATGAT 480
Qy 581 TCTTACTGAGTATTCGCGGCTTCTACCGGAGTGGCGATGCGAGAGTTGATTTCTTTTGA 640
Db 481 TCTTACTGAGTATTCGCGGCTTCTACCGGAGTGGCGATGCGAGAGTTGATTTCTTTTGA 540
Qy 641 ATTAAAGGCTGAGTTCGCGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 700
Db 541 ATTAAAGGCTGAGTTCGCGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 660
Qy 701 CAAACGAAAAAGAGATCCCTTCTCTGCTGCTGCGAGGTTTGAAGCTTCAACATTCCTT 760
Db 601 CAAACGAAAAAGAGATCCCTTCTCTGCTGCTGCGAGGTTTGAAGCTTCAACATTCCTT 660
Qy 761 GTGGTATTCGAGGTCAGGCGCTTATTTGGGCGGCTGTGATCTTTGGGCTGTGAGC 820
Db 661 GTGGTATTCGAGGTCAGGCGCTTATTTGGGCGGCTGTGATCTTTGGGCTGTGAGC 720
Qy 821 ATCCGATCTTCTCTTGGGAAAGGCTGCTAA 853
Db 721 ATCCGATCTTCTCTTGGGAAAGGCTGCTAA 753

RESULT 5

US-10-608-504-2
 ; Sequence 2, Application US/10608504
 ; Publication No. US20040014123A1
 ; GENERAL INFORMATION:
 ; APPLICANT: KENNERKNECHT, NICOLE
 ; APPLICANT: SAHM, HERMANN
 ; APPLICANT: EGGELING, LOTMAR
 ; APPLICANT: PFEFFERLE, WALTER
 ; TITLE OF INVENTION: NOCLEFTIDE SEQUENCE CODING FOR THE EXPORT OF
 ; TITLE OF INVENTION: BRANCHED CHAIN AMINO ACIDS, PROCESS FOR THE
 ; FILE REFERENCE: 7601/80525
 ; CURRENT APPLICATION NUMBER: US/10/608,504
 ; PRIOR FILING DATE: 2003-06-30
 ; PRIOR FILING DATE: 1999-12-23
 ; PRIOR APPLICATION NUMBER: DE 199 51 708.8
 ; PRIOR FILING DATE: 1999-10-27
 ; NUMBER OF SEQ ID NOS: 12
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 2
 ; LENGTH: 753
 ; TYPE: DNA
 ; ORGANISM: Corynebacterium glutamicum
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (1)..(753)
 ; OTHER INFORMATION: brnp
 ; FEATURE:
 ; OTHER INFORMATION: ATCC14752
 ; US-10-608-504-2

Query Match 59.2%; Score 753; DB 15; Length 753;
 Best Local Similarity 100.0%; Pred. No. 6,5e-245;

Matches 753; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 101 GTGCAAAAACGCAAGAGATTTCATTCAGACCTGAGAGGTGTGCGCATCCAGGACCCCTG 160
 DB 1 GTGCAAAAACGCAAGAGATTTCATTCAGACCTGAGAGGTGTGCGCATCCAGGACCCCTG 60
 QY 161 GAACGAGATGATTAAGTTATCGGGCTACGAAATCGGGCAAGGCTTAAACCTCCCTT 220
 DB 61 GAACGAGATGATTAAGTTATCGGGCTACGAAATCGGGCAAGGCTTAAACCTCCCTT 120
 QY 221 GCTGAGGTTTGGGAGTATGACCGATGCGTATGCGTTTGTCTCTTGTTATTCATATC 280
 DB 121 GCTGAGGTTTGGGAGTATGACCGATGCGTATGCGTTTGTCTCTTGTTATTCATATC 180
 QY 281 GGCTACGATGATGAGGAG 340
 DB 181 GGCTACGATGATGAGGAG 240
 QY 341 ATGCTGTCATGAGCCCTCGTTGTGGGCGAGAGCCCTGGGCGCATGCGCTACACCA 400
 DB 241 ATGCTGTCATGAGCCCTCGTTGTGGGCGAGAGCCCTGGGCGCATGCGCTACACCA 300
 QY 401 TTGCTGCTAATCTTCGCGACGATTTCTATGCGTTTCAATTCGCCCTGCAATGTGTCAA 460
 DB 301 TTGCTGCTAATCTTCGCGACGATTTCTATGCGTTTCAATTCGCCCTGCAATGTGTCAA 360
 QY 461 AACCCATGCGCGTTTCTATTCGTTTTCGCGTTATGAGAGAGAGAGAGAGAGAGAGAG 520
 DB 361 AACCCATGCGCGTTTCTATTCGTTTTCGCGTTATGAGAGAGAGAGAGAGAGAGAGAG 420
 QY 521 GCGGCGAG 580
 DB 421 GCGGCGAG 480
 QY 581 TCTTACTGGGATTTGCGGCGAGTCTACCGAGAGTGGCGAGTGGAGTGGATTCCTTTGAA 640
 DB 481 TCTTACTGGGATTTGCGGCGAGTCTACCGAGAGTGGCGAGTGGAGTGGATTCCTTTGAA 540

QY 641 ATTAGGAGCTGAGTTGCGCCCTTGTCTCTCTCTTTGTGACGCTGACTTGGATTCCTGC 700
 DB 541 ATTAGGAGCTGAGTTGCGCCCTTGTCTCTCTCTTTGTGACGCTGACTTGGATTCCTGC 600
 QY 701 CGAAGCAAAAAGAGATCCCTTCTCTGCTGCTCGAGGTTTGAAGTTCAACATTTGCTCT 760
 DB 601 CGAAGCAAAAAGAGATCCCTTCTCTGCTGCTCGAGGTTTGAAGTTCAACATTTGCTCT 660
 QY 761 GTGTAATTCAGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 820
 DB 661 GTGTAATTCAGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 720
 QY 821 ATCCGACTCTTCTTGGGAAAAGGCTGTAA 853
 DB 721 ATCCGACTCTTCTTGGGAAAAGGCTGTAA 753

RESULT 6

US-10-134-640-7/c
 ; Sequence 7, Application US/10134640
 ; Publication No. US20030017555A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Bathe, Brigitte
 ; APPLICANT: Kalinowski, Jörn
 ; APPLICANT: Puhler, Alfred
 ; APPLICANT: Mockel, Bettina
 ; APPLICANT: Pfefferle, Walter
 ; TITLE OF INVENTION: Nucleotide Sequences Coding for the lrp Gene
 ; FILE REFERENCE: 990109 BT
 ; CURRENT APPLICATION NUMBER: US/10/134,640
 ; CURRENT FILING DATE: 2002-04-30
 ; NUMBER OF SEQ ID NOS: 9
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 7
 ; LENGTH: 397
 ; TYPE: DNA
 ; ORGANISM: Corynebacterium glutamicum
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (1)..(397)
 ; OTHER INFORMATION: lrp part 1
 ; US-10-134-640-7

Query Match 28.8%; Score 365.6; DB 12; Length 397;

Best Local Similarity 98.9%; Pred. No. 3,3e-113;

Matches 368; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 9 AATGGAATCTAGCTTCAATATATGCAACAATAGCCTAGTGAAGGTGGCAAACTGGCAACA 68
 DB 376 AATGGAATCTAGCTTCAATATATGCAACAATAGCCTAGTGAAGGTGGCAAACTGGCAACA 317
 QY 69 AAATCAACCGGCAATGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 128
 DB 316 AAATCAACCGGCAATGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 257
 QY 129 GCGTGAAGTGTGCGCATCCAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 188
 DB 256 GCGTGAAGTGTGCGCATCCAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 197
 QY 189 ACGAAATCGGCAAGTCTAAAAAAGTCTCTGCTGAGGTTTGGGAGTATACCGATTG 248
 DB 196 ACGAAATCGGCAAGTCTAAAAAAGTCTCTGCTGAGGTTTGGGAGTATACCGATTG 137
 QY 249 GTATGAGTTTGTGCTCTGTTTATTCATATGAGGATACGAATGATGAGGAGAGAGAGAG 308
 DB 136 GTATGAGTTTGTGCTCTGTTTATTCATATGAGGATACGAATGATGAGGAGAGAGAGAG 77
 QY 309 TTTCGAGCTGATTTTTCGCGGCTCCACCGAAATGCTGATGCTGCTGTTGTGGGCG 368
 DB 76 TTTCGAGCTGATTTTTCGCGGCTCCACCGAAATGCTGATGCTGCTGTTGTGGGCG 17
 QY 369 CAGGCGCCCTGG 380

Db 16 CAGCCCCCGGG 5

RESULT 7

US-10-134-640-9/c

; Sequence 9, Application US/10134640

; Publication No. US20030017555A1

; GENERAL INFORMATION:

; APPLICANT: Baehr, Brigitte

; APPLICANT: Kalinowski, John

; APPLICANT: Puhler, Alfred

; APPLICANT: Mockel, Bettina

; APPLICANT: Pfefferle, Walter

; TITLE OF INVENTION: Nucleotide Sequences Coding for the lrp Gene

; FILE REFERENCE: 990109 BT

; CURRENT FILING DATE: 2002-04-30

; CURRENT APPLICATION NUMBER: US/10/134,640

; NUMBER OF SEQ ID NOS: 9

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 9

; LENGTH: 778

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; NAME/KEY: misc_feature

; LOCATION: (1)..(349)

; OTHER INFORMATION: upstream region of the lrp gene

; FEATURE:

; NAME/KEY: misc_feature

; LOCATION: (350)..(376)

; OTHER INFORMATION: 5'-region of the lrp-Gens

; FEATURE:

; NAME/KEY: misc_feature

; LOCATION: (377)..(397)

; OTHER INFORMATION: 21 bp extension

; FEATURE:

; NAME/KEY: misc_feature

; LOCATION: (398)..(433)

; OTHER INFORMATION: 3'-region of the lrp gene

; FEATURE:

; NAME/KEY: misc_feature

; LOCATION: (434)..(778)

; OTHER INFORMATION: downstream region of the lrp gene

; US-10-134-640-9

Query Match 28.8%; Score 365.6; DB 12; Length 778;

Best Local Similarity 98.9%; Pred. No. 5.2e-113;

Matches 368; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Db 9 AATGGAATCTAGCTTCATATATATGCAATAGGCTAGTGGGCAAACTGGCAACA 68

Db 376 AATGGAATCTAGCTTCATATATATGCAATAGGCTAGTGGGCAAACTGGCAACA 317

QY 69 AAATACCCGCAATTGTGATGATTTGATGTGCAAAAAACGCAAGATTCATTCAA 128

Db 316 AAATACCCGCAATTGTGATGATTTGATGTGCAAAAAACGCAAGATTCATTCAA 257

QY 129 GCCTGAGGTGTGCGCCATCCAGGAGCCCTGGAACGATGATAAAGTTATCGGCGCT 188

Db 129 GCCTGAGGTGTGCGCCATCCAGGAGCCCTGGAACGATGATAAAGTTATCGGCGCT 188

QY 256 GCGTGGAGGTGTGCGCCATCCAGGAGCCCTGGAACGATGATAAAGTTATCGGCGCT 197

Db 256 GCGTGGAGGTGTGCGCCATCCAGGAGCCCTGGAACGATGATAAAGTTATCGGCGCT 197

QY 189 ACGAATACGCAAGTCTAAAAAATCTCCCTGTGTCAGGTTTGGGATGTACCGAATTG 248

Db 196 ACGAATACGCAAGTCTAAAAAATCTCCCTGTGTCAGGTTTGGGATGTACCGAATTG 248

QY 249 GTATGCTGTTGCTCTTCTGTTTTCATACGCTTCGATGCTGGGCGAGCCCACTGT 308

Db 249 GTATGCTGTTGCTCTTCTGTTTTCATACGCTTCGATGCTGGGCGAGCCCACTGT 308

QY 136 GTATTTGGCTTGTCTCTTGTATTCATACGCTTCGATGCTGGGCGAGCCCACTGT 77

Db 136 GTATTTGGCTTGTCTCTTGTATTCATACGCTTCGATGCTGGGCGAGCCCACTGT 77

QY 309 TTTCGCGCTGATTTTGGCGGCTCCACCGAATAGTGTGATCGCCCTGCTTGTGGCG 368

Db 76 TTTCGCGCTGATTTTGGCGGCTCCACCGAATAGTGTGATCGCCCTGCTTGTGGCG 17

QY 369 CAGGCCCCCTGG 380

Db 16 CAGGCCCCCGGG 5

RESULT 8

US-09-738-626-290

; Sequence 290, Application US/09738626

; Publication No. US20020197605A1

; GENERAL INFORMATION:

; APPLICANT: NAKAGAWA, SATOSHI

; APPLICANT: MIZOGUCHI, HIROSHI

; APPLICANT: ANDO, SEIKO

; APPLICANT: HAYASHI, MIKIRO

; APPLICANT: OCHIAI, KEIKO

; APPLICANT: YOKOI, HARUHIKO

; APPLICANT: TATEISHI, MAKOTO

; APPLICANT: SENOH, AKIHIRO

; APPLICANT: IKEDA, MASATO

; APPLICANT: OZAKI, AKIO

; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES

; FILE REFERENCE: 249-125

; CURRENT APPLICATION NUMBER: US/09/738,626

; CURRENT FILING DATE: 2000-12-18

; PRIOR APPLICATION NUMBER: JP 99/377484

; PRIOR FILING DATE: 1999-12-16

; PRIOR APPLICATION NUMBER: JP 00/159162

; PRIOR FILING DATE: 2000-04-07

; PRIOR APPLICATION NUMBER: JP 00/280988

; PRIOR FILING DATE: 2000-08-03

; NUMBER OF SEQ ID NOS: 7059

; SOFTWARE: PatentIn ver. 3.0

; SEQ ID NO 290

; LENGTH: 324

; TYPE: DNA

; ORGANISM: Corynebacterium glutamicum

; US-09-738-626-290

Query Match 25.5%; Score 324; DB 9; Length 324;

Best Local Similarity 100.0%; Pred. No. 4.3e-99;

Matches 324; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 853 ATGACAACTGATTTCTCTGATTTCTCTGTTGTCGACAGTATGAGTCAATTTT 912

Db 1 ATGACAACTGATTTCTCTGATTTCTCTGTTGTCGACAGTATGAGTCAATTTT 60

QY 913 GCGCTCGGCGGCTTCGTTCTTATCTTAACTTAAAGCCCTAGCTGATCAATTTTGGGC 972

Db 61 GCGCTCGGCGGCTTCGTTCTTATCTTAACTTAAAGCCCTAGCTGATCAATTTTGGGC 120

QY 973 AAATGCGGATGTGAGTCCAGCAGAACTCTTGCATTTTGAACCGCATCACTTTCGC 1032

Db 121 AAATGCGGATGTGAGTCCAGCAGAACTCTTGCATTTTGAACCGCATCACTTTCGC 180

QY 1033 AGCAATGCGATGATCTGAACTCTTAACTTTCATTTGCTCATTTGCGATTGAGTG 1092

Db 181 AGCAATGCGATGATCTGAACTCTTAACTTTCATTTGCTCATTTGCGATTGAGTG 240

QY 1093 GTGGGCACTCTTCTTGGCGGTGCGACGACCTTGTGAGCGTTGGCGTGGACCATGCTT 1152

Db 241 GTGGGCACTCTTCTTGGCGGTGCGACGACCTTGTGAGCGTTGGCGTGGACCATGCTT 300

QY 1153 TTTGTTGACGTGTGAATCTTTTC 1176

Db 301 TTTGTTGACGTGTGAATCTTTTC 324

RESULT 9

US-10-608-504-4

; Sequence 4, Application US/10608504

; Publication No. US20040014123A1

; GENERAL INFORMATION:

; APPLICANT: KENNERKNECHT, NICOLE


```

; APPLICANT: SAKM, HERMANN
; APPLICANT: EGGELING, LOTMAR
; APPLICANT: PFEFFERLE, WALTER
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCE CODING FOR THE EXPORT OF
; TITLE OF INVENTION: BRANCHED CHAIN AMINO ACIDS, PROCESS FOR THE
; TITLE OF INVENTION: ISOLATION THEREOF AND USE THEREOF
; FILE REFERENCE: 7601/80525
; CURRENT APPLICATION NUMBER: US/10/608,504
; PRIOR FILING DATE: 2003-06-30
; PRIOR APPLICATION NUMBER: US 09/471,803
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: DE 199 S1 708.8
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 324
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(324)
; OTHER INFORMATION: brne
; FEATURE:
; OTHER INFORMATION: ATCC14752
; US-10-608-504-4

Query Match      25.5%; Score 324; DB 15; Length 324;
Best Local Similarity 100.0%; Pred. No. 4,3e-99;
Matches 324; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      853 ATGACACTGATTTCTCTCTGTAATTCCTCTGTTGTCGAGATGTCAGTCATTACTTTT 912
DB      1  ATGACAACTGATTTCTCTCGTAATTCCTCTGTTGTCGAGATGTCAGTCATTACTTTT 60

QY      913 GCGCTCCGGGCGGTCCTCCGTTTAACTCTTAAGCCCTACGAGATGCAATTGTCGAGC 972
DB      61  GCGCTCCGGGCGGTCCTCCGTTTAACTCTTAAGCCCTACGAGATGCAATTGTCGAGC 120

QY      973 AAATAGCGGATGATGTCGACGAGAAATCCTTGGCATTTTGAACCGCATCAACGTTTGC 1032
DB      121  AAAATAGCGGATGATGTCGACGAGAAATCCTTGGCATTTTGAACCGCATCAACGTTTGC 180

QY      1033 AGCATGCGATGATGTCGACGAGAAATCCTTGGCATTTTGAACCGCATCAACGTTTGC 1092
DB      181  AGCATGCGATGATGTCGACGAGAAATCCTTGGCATTTTGAACCGCATCAACGTTTGC 240

QY      1093 GTGGCGCATCTTCTGGCGGTCGACGACCTTGTGAGCGTTGCGCTGCGACCATCGTT 1152
DB      241  GTGGCGCATCTTCTGGCGGTCGACGACCTTGTGAGCGTTGCGCTGCGACCATCGTT 300

QY      1153 TTTGTTGACCTGCTGAATCTTTTC 1176
DB      301  TTTGTTGACCTGCTGAATCTTTTC 324

RESULT 10
US-10-134-640-1/c
; Sequence 1, Application US/10134640
; Publication No. US20030017555A1
; GENERAL INFORMATION:
; APPLICANT: Bathe, Brigitte
; APPLICANT: Kalinowski, Jorn
; APPLICANT: Puhler, Alfred
; APPLICANT: Mockel, Bettina
; APPLICANT: Pfeifferle, Walter
; TITLE OF INVENTION: Nucleotide Sequences Coding for the lrp Gene
; FILE REFERENCE: 990109 BT
; CURRENT APPLICATION NUMBER: US/10/134,640
; PRIOR FILING DATE: 2002-04-30
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1

```

```

; LENGTH: 715
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (151)..(612)
; OTHER INFORMATION: lrp-Gen
; FEATURE:
; NAME/KEY: -10 signal
; LOCATION: (88)..(93)
; OTHER INFORMATION:
; FEATURE:
; NAME/KEY: -35 signal
; LOCATION: (62)..(67)
; OTHER INFORMATION:
; US-10-134-640-1

Query Match      14.6%; Score 185; DB 12; Length 715;
Best Local Similarity 100.0%; Pred. No. 1,5e-51;
Matches 185; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1  GCGGATCAATGGAATCTAGCTCATATATTGCAACAATAGCCTAGTTGAGGTGCGAAAC 60
DB      185  GCGGATCAATGGAATCTAGCTCATATATTGCAACAATAGCCTAGTTGAGGTGCGAAAC 126

QY      61  TGGCAACAAACTAACCCGCAATTGTGTGATGATTTGATGTGCAAAAAACGCAAGAT 120
DB      125  TGGCAACAAACTAACCCGCAATTGTGTGATGATTTGATGTGCAAAAAACGCAAGAT 66

QY      121  TCATTCAAGCCTGAGAGTGTGCGCATTCGACAGAGCCCTGGAACGATGATAAGGTTA 180
DB      65  TCATTCAAGCCTGAGAGTGTGCGCATTCGACAGAGCCCTGGAACGATGATAAGGTTA 6

QY      181  TCGGC 185
DB      5  TCGGC 1

RESULT 11
US-09-738-626-3408
; Sequence 3408, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAOKO
; APPLICANT: SENO, AKIHITO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 3408
; LENGTH: 711
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
; US-09-738-626-3408

Query Match      10.0%; Score 127.4; DB 9; Length 711;
Best Local Similarity 51.5%; Pred. No. 6,1e-32;

```

```

1: SOFTWARE: PatentIn ver. 3.0
2: SEQ ID NO 1
3: LENGTH: 3309400
4: TYPE: DNA
5: ORGANISM: Corynebacterium glutamicum
6: US-09-738-626-1
7:
8: Query Match 10.0%; Score 127.4; DB 9; Length 3309400;
9: Best Local Similarity 15.5%; Pred. No. 1,4e-29;
10: Matches 316; Conservative 0; Mismatches 296; Indels 3; Gaps 1;
11:
12:
13:
14:
15:
16:
17:
18:
19:
20:
21:
22:
23:
24:
25:
26:
27:
28:
29:
30:
31:
32:
33:
34:
35:
36:
37:
38:
39:
40:
41:
42:
43:
44:
45:
46:
47:
48:
49:
50:
51:
52:
53:
54:
55:
56:
57:
58:
59:
60:
61:
62:
63:
64:
65:
66:
67:
68:
69:
70:
71:
72:
73:
74:
75:
76:
77:
78:
79:
80:
81:
82:
83:
84:
85:
86:
87:
88:
89:
90:
91:
92:
93:
94:
95:
96:
97:
98:
99:
100:
101:
102:
103:
104:
105:
106:
107:
108:
109:
110:
111:
112:
113:
114:
115:
116:
117:
118:
119:
120:
121:
122:
123:
124:
125:
126:
127:
128:
129:
130:
131:
132:
133:
134:
135:
136:
137:
138:
139:
140:
141:
142:
143:
144:
145:
146:
147:
148:
149:
150:
151:
152:
153:
154:
155:
156:
157:
158:
159:
160:
161:
162:
163:
164:
165:
166:
167:
168:
169:
170:
171:
172:
173:
174:
175:
176:
177:
178:
179:
180:
181:
182:
183:
184:
185:
186:
187:
188:
189:
190:
191:
192:
193:
194:
195:
196:
197:
198:
199:
200:
201:
202:
203:
204:
205:
206:
207:
208:
209:
210:
211:
212:
213:
214:
215:
216:
217:
218:
219:
220:
221:
222:
223:
224:
225:
226:
227:
228:
229:
230:
231:
232:
233:
234:
235:
236:
237:
238:
239:
240:
241:
242:
243:
244:
245:
246:
247:
248:
249:
250:
251:
252:
253:
254:
255:
256:
257:
258:
259:
260:
261:
262:
263:
264:
265:
266:
267:
268:
269:
270:
271:
272:
273:
274:
275:
276:
277:
278:
279:
280:
281:
282:
283:
284:
285:
286:
287:
288:
289:
290:
291:
292:
293:
294:
295:
296:
297:
298:
299:
300:
301:
302:
303:
304:
305:
306:
307:
308:
309:
310:
311:
312:
313:
314:
315:
316:
317:
318:
319:
320:
321:
322:
323:
324:
325:
326:
327:
328:
329:
330:
331:
332:
333:
334:
335:
336:
337:
338:
339:
340:
341:
342:
343:
344:
345:
346:
347:
348:
349:
350:
351:
352:
353:
354:
355:
356:
357:
358:
359:
360:
361:
362:
363:
364:
365:
366:
367:
368:
369:
370:
371:
372:
373:
374:
375:
376:
377:
378:
379:
380:
381:
382:
383:
384:
385:
386:
387:
388:
389:
390:
391:
392:
393:
394:
395:
396:
397:
398:
399:
400:
401:
402:
403:
404:
405:
406:
407:
408:
409:
410:
411:
412:
413:
414:
415:
416:
417:
418:
419:
420:
421:
422:
423:
424:
425:
426:
427:
428:
429:
430:
431:
432:
433:
434:
435:
436:
437:
438:
439:
440:
441:
442:
443:
444:
445:
446:
447:
448:
449:
450:
451:
452:
453:
454:
455:
456:
457:
458:
459:
460:
461:
462:
463:
464:
465:
466:
467:
468:
469:
470:
471:
472:
473:
474:
475:
476:
477:
478:
479:
480:
481:
482:
483:
484:
485:
486:
487:
488:
489:
490:
491:
492:
493:
494:
495:
496:
497:
498:
499:
500:
501:
502:
503:
504:
505:
506:
507:
508:
509:
510:
511:
512:
513:
514:
515:
516:
517:
518:
519:
520:
521:
522:
523:
524:
525:
526:
527:
528:
529:
530:
531:
532:
533:
534:
535:
536:
537:
538:
539:
540:
541:
542:
543:
544:
545:
546:
547:
548:
549:
550:
551:
552:
553:
554:
555:
556:
557:
558:
559:
560:
561:
562:
563:
564:
565:
566:
567:
568:
569:
570:
571:
572:
573:
574:
575:
576:
577:
578:
579:
580:
581:
582:
583:
584:
585:
586:
587:
588:
589:
590:
591:
592:
593:
594:
595:
596:
597:
598:
599:
600:
601:
602:
603:
604:
605:
606:
607:
608:
609:
610:
611:
612:
613:
614:
615:
616:
617:
618:
619:
620:
621:
622:
623:
624:
625:
626:
627:
628:
629:
630:
631:
632:
633:
634:
635:
636:
637:
638:
639:
640:
641:
642:
643:
644:
645:
646:
647:
648:
649:
650:
651:
652:
653:
654:
655:
656:
657:
658:
659:
660:
661:
662:
663:
664:
665:
666:
667:
668:
669:
670:
671:
672:
673:
674:
675:
676:
677:
678:
679:
680:
681:
682:
683:
684:
685:
686:
687:
688:
689:
690:
691:
692:
693:
694:
695:
696:
697:
698:
699:
700:
701:
702:
703:
704:
705:
706:
707:
708:
709:
710:
711:
712:
713:
714:
715:
716:
717:
718:
719:
720:
721:
722:
723:
724:
725:
726:
727:
728:
729:
730:
731:
732:
733:
734:
735:
736:
737:
738:
739:
740:
741:
742:
743:
744:
745:
746:
747:
748:
749:
750:
751:
752:
753:
754:
755:
756:
757:
758:
759:
760:
761:
762:
763:
764:
765:
766:
767:
768:
769:
770:
771:
772:
773:
774:
775:
776:
777:
778:
779:
780:
781:
782:
783:
784:
785:
786:
787:
788:
789:
790:
791:
792:
793:
794:
795:
796:
797:
798:
799:
800:
801:
802:
803:
804:
805:
806:
807:
808:
809:
810:
811:
812:
813:
814
```

PRIOR FILING DATE: 2001-05-30
PRIOR APPLICATION NUMBER: JP 2001-272697
PRIOR FILING DATE: 2001-08-02
NUMBER OF SEQ ID NOS: 15109
SEQ ID NO 6067
LENGTH: 594
TYPE: DNA
ORGANISM: Streptomyces avermitilis
FEATURE:
NAME/KEY: CDS
LOCATION: (1)..(594)
US-10-156-761-6067

Query Match 3.4%; Score 43; DB 14; Length 594;
Best Local Similarity 50.7%; Pred. No. 0.0029;
Matches 103; Conservative 0; Mismatches 100; Indels 0; Gaps 0;

QY 324 TCGGGGCTCCACGAAATGCTGTCATCGCCCTCGTTGTGGGCGCAGCGCCCTGGGCG 383
DB 203 TCACCGGCTGGCCATCGCGCAGGTGCTCGGCATGTTGTGGCAGCGCGCTCGGCTGGG 262
QY 384 CCATCGCGCTCAACCAATGCTGTGTAATTCGCCACGATTTATGCGTTTCATTCC 443
DB 263 GCAACGTGCCACGACGATCTCGGCGATGCTGCTGCTCTTCTTGGGCTACTGCTCA 322
QY 444 CGCTGCATGTGTCAAAACCCCATTTGCCGTTCTATTGCTTTCGCGCTTATGACG 503
DB 323 CGTTGGCGGAGTGTGTGAAGCGCGGTGTCGCTTCGCGACCGCCTTGGGGGTCCGCTCG 382
QY 504 AAGCTACGACATCACTCGCGCC 526
DB 383 CCGCGACACCCCTGTCATCGCC 405

RESULT 14

US-10-156-761-1/c
Sequence 1, Application US/10156761
Publication No. US20030119018A1
GENERAL INFORMATION:
APPLICANT: OMURA, SATOSHI
APPLICANT: IKEDA, HARUO
APPLICANT: ISHIKAWA, JUN
APPLICANT: HORIKAWA, HIROSHI
APPLICANT: SHIBA, TADAYOSHI
APPLICANT: SAKAKI, YOSHIYUKI
APPLICANT: HATTORI, MASAHIRA
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-262
CURRENT APPLICATION NUMBER: US/10/156,761
CURRENT FILING DATE: 2002-05-29
PRIOR APPLICATION NUMBER: JP 2001-204089
PRIOR FILING DATE: 2001-05-30
PRIOR APPLICATION NUMBER: JP 2001-272697
PRIOR FILING DATE: 2001-08-02
NUMBER OF SEQ ID NOS: 15109
SEQ ID NO 1
LENGTH: 9025608
TYPE: DNA
ORGANISM: Streptomyces avermitilis
FEATURE:
NAME/KEY: misc feature
LOCATION: (4187715)
OTHER INFORMATION: a, t, c, g, other or unknown
US-10-156-761-1

Query Match 3.4%; Score 43; DB 14; Length 9025608;
Best Local Similarity 50.7%; Pred. No. 1.4;
Matches 103; Conservative 0; Mismatches 100; Indels 0; Gaps 0;

QY 324 TCGGGGCTCCACGAAATGCTGTCATCGCCCTCGTTGTGGGCGCAGCGCCCTGGGCG 383
DB 7327430 TCACCGGCTGGCCATCGCGCAGGTGCTCGGCATGTTGTGGCAGCGCGCTCGGCTGGG 7327371

QY 384 CCATCGCGCTCAACCAATGCTGTGTAATTCGCCACGATTTATGCGTTTCATTCC 443
DB 7327370 GCAACGTGCCACGACGATCTCGGCGATGCTGCTGCTTCTTTCGCGCTACTGCTCA 7327311
QY 444 CGCTGCATGTGTCAAAACCCCATTTGCCGTTCTATTGCTTTCGCGCTTATGACG 503
DB 7327310 CGTTGGCGGAGTGTGTGAAGCGCGGTGTCGCTTCGCGACCGCCTTGGGGGTCCGCTCG 7327251
QY 504 AAGCTACGACATCACTCGCGCC 526
DB 7327250 CCGCGACACCCCTGTCATCGCC 7327228

RESULT 15

US-10-184-644-80/c
Sequence 80, Application US/10184644
Publication No. US20030044930A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Chen, Jian
APPLICANT: Desnoyers, Luc
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Smith, Victoria
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3430R1C227
CURRENT APPLICATION NUMBER: US/10/184,644
CURRENT FILING DATE: 2002-06-28
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 612
SEQ ID NO 80
LENGTH: 351
TYPE: PRT
ORGANISM: Homo Sapien
US-10-184-644-80

Query Match 3.3%; Score 41.4; DB 14; Length 351;
Best Local Similarity 13.4%; Pred. No. 0.0073;
Matches 45; Conservative 103; Mismatches 187; Indels 0; Gaps 0;

QY 161 GAACCAATGATTAAGGTTATCGCGCTACGAATCGCGCAGGTCTAAACCTCCCTT 220
DB 350 GAD..B.Y.G.YGMB.GGMD.MSGV.T..B.TAD.CYCGG...GCH.TMSANGAICYA 291
QY 221 GCTGCAAGTTGGGCAATGACCCGATTTGCTGTTGCTCTTGGTTATCAATAC 280
DB 290 C.BESYGH..G.H.S...RB.R.HAD.M..TNY.B..TYSCEBY..B.ATCH.K.M.TDC 231
QY 281 GGCTACGATGTGGGAGCCCACTGTTTTCGCGCTGATTTTCGCGGCTCCACGAA 340
DB 230 B.Y...BTYGMAB.GSCBDDMG.BYKDA..TMYCT.NND.NK.TYSSSTCMY.YBH 171
QY 341 ATGCTGTCATCGCCCTGTTGTGGGCGCAGCGCCCTGAGCCCATCGGCTCACCA 400
DB 170 B.S.S.TBCRT.NHSGSW..C..SBCDBAYHCGM...CSGTVAAT.TBCTSS.TRB. 111
QY 401 TTGCTGTGAATCTCCGCAAGTATTGATCGCTTTTCATTCGCGCTGATGTGTCAAA 460
DB 110 TT.YAC.ABYVC.B..SH.ASGMYNY..R.G...T..TCYBSHKB..NYMK.B.Y. 51
QY 461 AACCACATCGCCCTTCTATTCGTTTCGCGCT 495
DB 50 ..CM.M.GA...S.CBSS.MT.R.MNSTTTS.BT 16

Search completed: April 17, 2004, 21:00:17
Job time : 647.884 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 19, 2004, 10:57:18 ; Search time 4560.51 Seconds

(without alignments)
10226.291 Million cell updates/sec

Title: US-10-608-504-1_COPY_101_1176

Perfect score: 1076

Sequence: 1 gtgcaaaaacgcgaagatc.....ctggactggtgaatcttcc 1076

Scoring table: IDENTITY_NUC

Gapop 10.0 ; Gapext 1.0

Searched: 3470272 seqs, 2167151695 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : GenEmbl.*

1: gb_ba.*

2: gb_hcg.*

3: gb_in.*

4: gb_ov.*

5: gb_ov.*

6: gb_ov.*

7: gb_ov.*

8: gb_ov.*

9: gb_ov.*

10: gb_ov.*

11: gb_ov.*

12: gb_ov.*

13: gb_ov.*

14: gb_ov.*

15: gb_ov.*

16: gb_ov.*

17: gb_ov.*

18: gb_ov.*

19: gb_ov.*

20: gb_ov.*

21: gb_ov.*

22: gb_ov.*

23: gb_ov.*

24: gb_ov.*

25: gb_ov.*

26: gb_ov.*

27: gb_ov.*

28: gb_ov.*

29: gb_ov.*

30: gb_ov.*

31: gb_ov.*

32: gb_ov.*

33: gb_ov.*

34: gb_ov.*

35: gb_ov.*

36: gb_ov.*

37: gb_ov.*

38: gb_ov.*

39: gb_ov.*

40: gb_ov.*

41: gb_ov.*

score greater than or equal to the score being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1076	100.0	1271	6	AR391953 Sequence
2	1076	100.0	1271	6	AR391956 Sequence
3	1076	100.0	1271	6	AR391957 Sequence
4	1076	100.0	1271	6	AR391958 Sequence
5	1076	100.0	1271	6	AR391959 Sequence
6	1076	100.0	1271	6	AR391960 Sequence
7	1076	100.0	1271	6	AR391961 Sequence
8	1076	100.0	1271	6	AR391962 Sequence
9	1076	100.0	1271	6	AR391963 Sequence
10	1076	100.0	1271	6	AR391964 Sequence
11	1076	100.0	1271	6	AR391965 Sequence
12	1076	100.0	1271	6	AR391966 Sequence
13	1076	100.0	1271	6	AR391967 Sequence
14	1076	100.0	1271	6	AR391968 Sequence
15	1076	100.0	1271	6	AR391969 Sequence
16	1076	100.0	1271	6	AR391970 Sequence
17	1076	100.0	1271	6	AR391971 Sequence
18	1076	100.0	1271	6	AR391972 Sequence
19	1076	100.0	1271	6	AR391973 Sequence
20	1076	100.0	1271	6	AR391974 Sequence
21	1076	100.0	1271	6	AR391975 Sequence
22	1076	100.0	1271	6	AR391976 Sequence
23	1076	100.0	1271	6	AR391977 Sequence
24	1076	100.0	1271	6	AR391978 Sequence
25	1076	100.0	1271	6	AR391979 Sequence
26	1076	100.0	1271	6	AR391980 Sequence
27	1076	100.0	1271	6	AR391981 Sequence
28	1076	100.0	1271	6	AR391982 Sequence
29	1076	100.0	1271	6	AR391983 Sequence
30	1076	100.0	1271	6	AR391984 Sequence
31	1076	100.0	1271	6	AR391985 Sequence
32	1076	100.0	1271	6	AR391986 Sequence
33	1076	100.0	1271	6	AR391987 Sequence
34	1076	100.0	1271	6	AR391988 Sequence
35	1076	100.0	1271	6	AR391989 Sequence
36	1076	100.0	1271	6	AR391990 Sequence
37	1076	100.0	1271	6	AR391991 Sequence
38	1076	100.0	1271	6	AR391992 Sequence
39	1076	100.0	1271	6	AR391993 Sequence
40	1076	100.0	1271	6	AR391994 Sequence
41	1076	100.0	1271	6	AR391995 Sequence

ALIGNMENTS

RESULT 1

LOCUS AR391953

DEFINITION Sequence 1 from patent US 6613545.

ACCESSION AR391953

VERSION AR391953.1 GI:40115724

KEYWORDS

SOURCE Unknown.

ORGANISM Unknown.

REFERENCE 1 (bases 1 to 1271)

AUTHORS Kemeriknecht, N., Sahm, H., Eggeling, L. and Pfeifferle, W.

TITLE Nucleotide sequences coding for the export of branched chain amino acids, process for the isolation thereof and use thereof

JOURNAL Patent: US 6613545-A 1 02-SEP-2003;

FEATURES

source

Location/Qualifiers
1..1271
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN

Query Match 100.0%; Score 1076; DB 6; Length 1271;
Best Local Similarity 100.0%; Pred. No. 8.3e-275;
Matches 1076; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 GTGCAAAAAACGCAAGATTCATTCAAGCTGAGGTCTCCCATCCAGGACGCGCTG 60
101 GTGCAAAAAACGCAAGATTCATTCAAGCTGAGGTCTCCCATCCAGGACGCGCTG 160
61 GAAACGATGATTAAGGTTATCGGCGCTACGAAATGCGGCAAGGTTAAAACTCCCTT 120
161 GAAACGATGATTAAGGTTATCGGCGCTACGAAATGCGGCAAGGTTAAAACTCCCTT 220
121 GCTGACGAGTTGGGCAATGATACCGGATTTGGATTCCTTTGTTATTCATATAC 180
221 GCTGACGAGTTGGGCAATGATACCGGATTTGGATTCCTTTGTTATTCATATAC 280
181 GGTACGATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 240
281 GGTACGATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 340
241 ATGCTGATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 300
341 ATGCTGATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 400
301 TTGCTGTGAATCTTCGCGCAGTATTCATGCGTTTCAATCCGCTGATGTGCTGAA 360
401 TTGCTGTGAATCTTCGCGCAGTATTCATGCGTTTCAATCCGCTGATGTGCTGAA 460
361 AACCCATGCGCGGTTCTATGCGTTTCCGCTTATGAGGAACTTACGAGTCACT 420
461 AACCCATGCGCGGTTCTATGCGTTTCCGCTTATGAGGAACTTACGAGTCACT 520
421 GCGGCGAGGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 480
521 GCGGCGAGGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 580
481 TCTTACGAGGATTCGCGGCTTCAACGAGTGGGCAATGCGGATTCCTTTGAA 540
581 TCTTACGAGGATTCGCGGCTTCAACGAGTGGGCAATGCGGATTCCTTTGAA 640
541 ATTAAGGAGCTCGAGTTCGCGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTG 600
641 ATTAAGGAGCTCGAGTTCGCGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTG 700
601 GAAACGAAAAACGATTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 660
701 GAAACGAAAAACGATTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 760
661 GTGGTATTCAGGATTCGCGGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTG 720
761 GTGGTATTCAGGATTCGCGGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTG 820
721 ATCGGATTCCT 780
821 ATCGGATTCCT 880
781 TTGTTGCGAGTATGAGGAGTCACTTCTTTGCGCTCCGCGGCTTCGTTCTATATC 840
881 TTGTTGCGAGTATGAGGAGTCACTTCTTTGCGCTCCGCGGCTTCGTTCTATATC 940
841 TTAAGCCCTTACGTAATCAATTTGCGGCAAAATGCGAGTGTGATCCGACGAGAA 900
941 TTAAGCCCTTACGTAATCAATTTGCGGCAAAATGCGAGTGTGATCCGACGAGAA 1000
901 TCCCTGCGATTTTACCGGATCAAGCTTGGGCAAGTGGATGATGATGATGATGATGAT 960
1001 TCCCTGCGATTTTACCGGATCAAGCTTGGGCAAGTGGATGATGATGATGATGATGAT 1060

QY 961 CCTTGTGATTCATTCGCGCTTGGATTAAGTGTGGGCGATCTTCTGGGAGTCAAGCA 1020
DB 1061 CCTTGTGATTCATTCGCGCTTGGATTAAGTGTGGGCGATCTTCTGGGAGTCAAGCA 1120
QY 1021 CCTTGTGATTCATTCGCGCTTGGATTAAGTGTGGGCGATCTTCTGGGAGTCAAGCA 1076
DB 1121 CCTTGTGATTCATTCGCGCTTGGATTAAGTGTGGGCGATCTTCTGGGAGTCAAGCA 1176

RESULT 2
AR391956 1271 bp DNA linear PAT 18-DEC-2003
LOCUS AR391956
DEFINITION Sequence 6 from patent US 6613545.
ACCESSION AR391956
VERSION AR391956.1 GI:40115727
KEYWORDS
SOURCE
ORGANISM
Unknown.
REFERENCE
1 (bases 1 to 1271)
AUTHORS Kennenkrech, N., Sahm, H., Eggeling, L. and Pfeifferle, W.
TITLE Nucleotide sequences coding for the export of branched chain amino acids, process for the isolation thereof and use thereof
JOURNAL Patent: US 6613545-A 6 02-SEP-2003;
FEATURES
Location/Qualifiers
source
1..1271
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN

Query Match 100.0%; Score 1076; DB 6; Length 1271;
Best Local Similarity 100.0%; Pred. No. 8.3e-275;
Matches 1076; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 GTGCAAAAAACGCAAGATTCATTCAAGCTGAGGTCTCCCATCCAGGACGCGCTG 60
101 GTGCAAAAAACGCAAGATTCATTCAAGCTGAGGTCTCCCATCCAGGACGCGCTG 160
61 GAAACGATGATTAAGGTTATCGGCGCTACGAAATGCGGCAAGGTTAAAACTCCCTT 120
161 GAAACGATGATTAAGGTTATCGGCGCTACGAAATGCGGCAAGGTTAAAACTCCCTT 220
121 GCTGACGAGTTGGGCAATGATACCGGATTTGGATTCCTTTGTTATTCATATAC 180
221 GCTGACGAGTTGGGCAATGATACCGGATTTGGATTCCTTTGTTATTCATATAC 280
181 GGTACGATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 240
281 GGTACGATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 340
241 ATGCTGATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 300
341 ATGCTGATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 400
301 TTGCTGTGAATCTTCGCGCAGTATTCATGCGTTTCAATCCGCTGATGTGCTGAA 360
401 TTGCTGTGAATCTTCGCGCAGTATTCATGCGTTTCAATCCGCTGATGTGCTGAA 460
361 AACCCATGCGCGGTTCTATGCGTTTCCGCTTATGAGGAACTTACGAGTCACT 420
461 AACCCATGCGCGGTTCTATGCGTTTCCGCTTATGAGGAACTTACGAGTCACT 520
421 GCGGCGAGGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 480
521 GCGGCGAGGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 580
481 TCTTACGAGGATTCGCGGCTTCAACGAGTGGGCAATGCGGATTCCTTTGAA 540
581 TCTTACGAGGATTCGCGGCTTCAACGAGTGGGCAATGCGGATTCCTTTGAA 640
541 ATTAAGGAGCTCGAGTTCGCGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTG 600

Db 641 ATTAAGGGCCCTGAGTTGCGCCCTTGTCTCTCTTTGTGCAAGCTGACTTTGATTCCTGC 700
 QY 601 CGAAGCAAAAAGCAGATCCCTTCTCTGCTGCTGCGAGGTTTGAAGCTTACCAATTCCTT 660
 Db 701 CGAAGCAAAAAGCAGATCCCTTCTCTGCTGCTGCGAGGTTTGAAGCTTACCAATTCCTT 760
 QY 661 GTGGTAATTCAGAGTCAAGGCGCCCTATTTGGCGGCGTGCAGTACTTCTGGGGCTGTGACC 720
 Db 761 GTGGTAATTCAGAGTCAAGGCGCCCTATTTGGCGGCGTGCAGTACTTCTGGGGCTGTGACC 820
 QY 721 ATCCGTAATTCCTTCTTGGGAAAAGGCTGCTAAATGACAGTATTTCTCTGATTTCTCC 780
 Db 821 ATCCGTAATTCCTTCTTGGGAAAAGGCTGCTAAATGACAGTATTTCTCTGATTTCTCC 880
 QY 781 TTGTTGTGCGAGTATGTCAGTACTTACTTTTGGGCTCCGGGCGGTTCCGTTCTATCC 840
 Db 881 TTGTTGTGCGAGTATGTCAGTACTTACTTTTGGGCTCCGGGCGGTTCCGTTCTATCC 940
 QY 841 TTAAGCCCTACGTAATCAATTTGTGGCAAAATGGCGATGTGATGCGCAGAGAA 900
 Db 941 TTAAGCCCTACGTAATCAATTTGTGGCAAAATGGCGATGTGATGCGCAGAGAA 1000
 QY 901 TCCTTGGCATTTTGAACGCGATCAACGTTTGGCAGCAATGCGATGATCTGAAGCTTAA 960
 Db 1001 TCCTTGGCATTTTGAACGCGATCAACGTTTGGCAGCAATGCGATGATCTGAAGCTTAA 1060
 QY 961 CCTTGGTCTCATTCGCGGCTGCGATTAACAGTGTGGCGCATCTTCTTGGCGGTGACGCA 1020
 Db 1061 CCTTGGTCTCATTCGCGGCTGCGATTAACAGTGTGGCGCATCTTCTTGGCGGTGACGCA 1120
 QY 1021 CCTTGGTGGAGCGTTGGCGGCTGCGACCATGTTTTTGTGACGTGTAATCTTTTC 1076
 Db 1121 CCTTGGTGGAGCGTTGGCGGCTGCGACCATGTTTTTGTGACGTGTAATCTTTTC 1176

RESULT 3
 AX137709 1271 bp DNA linear PAT 30-MAY-2001
 LOCUS AX137709
 DEFINITION Sequence 1 from Patent EP1096010.
 ACCESSION AX137709
 VERSION AX137709.1 GI:14273886
 KEYWORDS
 SOURCE Corynebacterium glutamicum
 ORGANISM Corynebacterium glutamicum
 Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 Corynebacteriaceae; Corynebacterium.
 REFERENCE 1
 Kennerknecht, N., Eggeling, L., Sahm, H. and Pfefferle, W.
 Nucleotide sequences coding for branched-chain amino acids export
 proteins, method for isolating them and their use
 Patent: EP 1096010-A 1 02-MAY-2001;
 Degussa AG (DE) ; FORSCHUNGSZENTRUM JUELICH GMBH (DE)
 JOURNAL
 FEATURES
 source Location/Qualifiers
 1..1271
 /organism="Corynebacterium glutamicum"
 /mol_type="unassigned DNA"
 /db_xref="taxon:1718"
 /note="ATCC4752"
 101..856
 /note="unnamed protein product; brnF"
 /codon_start=1
 /transl_table=1
 /protein_id="CAC39923.1"
 /db_xref="GI:14273887"
 /db_xref="REMBL:CA39923.1"
 /translation="MKTQELHSHLEVPSPAIEPDDKGRYRIAGLTKTSIAGL
 GMYPIGAFGLVIOGYEWMAAPLPSGLIFAGSTEMVLAVVGAAPLGAIALTTLL
 VNFVFAFSPPLVNVNPIARFVSFALIDEAVAVTAAPAPGASWRLISMOIAH
 SYVFGGLGVAIALIPEIKGFALCSLFTVLTDSCKTKQIPILLAGSFTI
 ALVPIGQALPAALMIFGLLTIRYFISKAH"
 853..1179
 /note="unnamed protein product; brnE"
 /codon_start=1

CDS

/transl_table=1
 /protein_id="CAC39923.1"
 /db_xref="GI:14273887"
 /db_xref="REMBL:CA39923.1"
 /translation="MKTQELHSHLEVPSPAIEPDDKGRYRIAGLTKTSIAGL
 GMYPIGAFGLVIOGYEWMAAPLPSGLIFAGSTEMVLAVVGAAPLGAIALTTLL
 VNFVFAFSPPLVNVNPIARFVSFALIDEAVAVTAAPAPGASWRLISMOIAH
 SYVFGGLGVAIALIPEIKGFALCSLFTVLTDSCKTKQIPILLAGSFTI
 ALVPIGQALPAALMIFGLLTIRYFISKAH"
 GVNF"

Query Match 100.0%; Score 1076; DB 6; Length 1271;
 Best Local Similarity 100.0%; Pred. No. 8.3e-275;
 Matches 1076; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTGCAAAAAGCAGAGATTCATTCAAGCTCGAGAGTGTGCGCATCCAGGACGCGCTG 60
 Db 101 GTGCAAAAAGCAGAGATTCATTCAAGCTCGAGAGTGTGCGCATCCAGGACGCGCTG 160
 QY 61 GAACGATGATPAAGGTTATCGGCGCTACGAATCGGCAAGTCTAAAACTTCCTT 120
 Db 161 GAACGATGATPAAGGTTATCGGCGCTACGAATCGGCAAGTCTAAAACTTCCTT 220
 QY 121 GCTGCAAGTTGGGCAATGACCGGATGGATTTGGGTTGCTCTGTGTTATTCATAC 180
 Db 221 GCTGCAAGTTGGGCAATGACCGGATGGATTTGGGTTGCTCTGTGTTATTCATAC 280
 QY 181 GGCTACGAATGATGAGGACGCCCATCTGTTTCGGGCTGATTTTCGGGCGTCCACGAA 240
 Db 281 GGCTACGAATGATGAGGACGCCCATCTGTTTCGGGCTGATTTTCGGGCGTCCACGAA 340
 QY 241 ATGCTGATATGACCTCTGTTGTGGGCGGAGGCCCTGGGCGCATGCGGCTCACGAA 300
 Db 341 ATGCTGATATGACCTCTGTTGTGGGCGGAGGCCCTGGGCGCATGCGGCTCACGAA 400
 QY 301 TTGCTGTGAACCTCCGACAGATTTCTATGCGTTTTCATTCCTCCGTCATGTGTCAA 360
 Db 401 TTGCTGTGAACCTCCGACAGATTTCTATGCGTTTTCATTCCTCCGTCATGTGTCAA 460
 QY 361 AACCCATGCGCCGTTTCTATTCGCTTTTCGGGCTTATGAGAGAACCTGACGACAT 420
 Db 461 AACCCATGCGCCGTTTCTATTCGCTTTTCGGGCTTATGAGAGAACCTGACGACAT 520
 QY 421 GCGGACAGGCGCGCAGGCTGTGCGGCGTGGCAGCTTATCTCATGCAATAGCGTTTCA 480
 Db 521 GCGGACAGGCGCGCAGGCTGTGCGGCGTGGCAGCTTATCTCATGCAATAGCGTTTCA 580
 QY 481 TCCTACTGAGTATTCGCGGCTGTCACCGGAGTGGCGATGCGAGAGTTGATTCCTTTGAA 540
 Db 581 TCCTACTGAGTATTCGCGGCTGTCACCGGAGTGGCGATGCGAGAGTTGATTCCTTTGAA 640
 QY 541 ATTAAGGCGCTGAGTTCGCCCTTGTGCTCTCTCTTTGACGCTGATTCGATTCCTGC 600
 Db 641 ATTAAGGCGCTGAGTTCGCCCTTGTGCTCTCTCTTTGACGCTGATTCGATTCCTGC 700
 QY 601 CGAAGCAAAAAGCAGATCCCTTCTCTGCTGCTGCGAGTTTGAAGCTTACCAATTCCTT 660
 Db 701 CGAAGCAAAAAGCAGATCCCTTCTCTGCTGCTGCGAGTTTGAAGCTTACCAATTCCTT 760
 QY 661 GTGGTAATTCAGAGTCAAGGCGCCATTTGGGCGGCTGCTATCTTGGGCTGTGACC 720
 Db 761 GTGGTAATTCAGAGTCAAGGCGCCATTTGGGCGGCTGCTATCTTGGGCTGTGACC 820
 QY 721 ATCCGTAATTCCTTCTTGGGAAAAGGCTGCTAAATGACAACTGATTTCTCTGATTTCTCC 780
 Db 821 ATCCGTAATTCCTTCTTGGGAAAAGGCTGCTAAATGACAACTGATTTCTCTGATTTCTCC 880
 QY 781 TTGTTGTGCGAGTATGTCAGTACTTACTTTTGGGCTCCGGGCGGTTCCGTTCTTATCC 840
 Db 881 TTGTTGTGCGAGTATGTCAGTACTTACTTTTGGGCTCCGGGCGGTTCCGTTCTTATCC 940
 QY 841 TTAAGCCCTACGTAATCAATTTGTGGCAAAATGGCGATGTGATGCCAGAGAA 900
 Db 941 TTAAGCCCTACGTAATCAATTTGTGGCAAAATGGCGATGTGATGCCAGAGAA 1000

QY 901 TCCTGGCATTTGACGCGATCAAGTTTCGACGAATGCGATAGATCTGAAGCTCTAA 960
DB 1001 TCCTGGCATTTGACGCGATCAAGTTTCGACGAATGCGATAGATCTGAAGCTCTAA 1060
QY 961 CCTTTGGTCTCATTTGCGGCTTGGCATTTACAGTGTGGGCGCATCTTTGGCGGTGACGCA 1020
DB 1061 CCTTTGGTCTCATTTGCGGCTTGGCATTTACAGTGTGGGCGCATCTTTGGCGGTGACGCA 1120
QY 1021 CCTTTGAGCGGTGGGCGCTGACGATCTTTTGTGTGACTGGTGAATCTTTTC 1076
DB 1121 CCTTTGAGCGGTGGGCGCTGACGATCTTTTGTGTGACTGGTGAATCTTTTC 1176

RESULT 4
AX137714 1271 bp DNA linear PAT 30-MAY-2001
LOCUS Sequence 6 from Patent EP1096010.
DEFINITION AX137714
ACCESSION AX137714
VERSION AX137714.1 GI:14273893
KEYWORDS
SOURCE Corynebacterium glutamicum
ORGANISM Corynebacterium glutamicum
Bacteria; Actinobacteria; Actinomycetales;
Corynebacteriaceae; Corynebacteriaceae; Corynebacterium.
REFERENCE 1
AUTHORS Kemernechtz, N., Eggeling, L., Salm, H. and Pfeifferle, W.
TITLE Nucleotide sequences coding for branched-chain amino acids export proteins, method for isolating them and their use
JOURNAL Patent: EP 1096010-A 6 02-MAY-2001;
Degussa AG (DE) ; FORSCHUNGSZENTRUM JUELICH GMBH (DE)
FEATURES
source Location/Qualifiers
1.1271
/organism="Corynebacterium glutamicum"
/mol_type="unassigned DNA"
/db_xref="taxon:1718"
/note="ATCC13032"
101.856
/note="unnamed protein product; brnF"
/codon_start=1
/transl_table=1
/protein_id="CAC39927.1"
/db_xref="GI:14273894"
/db_xref="RMBTREMBL:CAC39927"
/translation="MOQTQEIHSLSLEVPKRALEPDDKGRYREIAQGLTSLAAGL
GVYPIAGFLVIOGYEMWAAPIFSGLTIFAGSTEMVILVVGAAPIGLATITLL
VNEHVFAPSPFLVKNPITAPKSPVAFIDBAVAVTAAPADGMSAMRLISQIAH
SYWVREGATGVAAIAELIPETIKGEPALCSLPVLTITLDSCTKQIISLLAGSFTI
ALVITPGQALPAAALITFLGLTIRIFLGKAK"
853.1179
/note="unnamed protein product; brnE"
/codon_start=1
/transl_table=1
/protein_id="CAC39928.1"
/db_xref="GI:14273895"
/db_xref="RMBTREMBL:CAC39928"
/translation="WTDPSCLLVAVCAVITPAAVAPVPIILKPIRESFVGKMAN
WMPAGILALITLSTFRSNAIDIKLITFLGILAVALTVAHLIGRRITLISVAGTIVFV
GLWLPF"

ORIGIN
Query Match 100.0% Score 1076; DB 6; Length 1271;
Best Local Similarity 100.0%; Pred. No. 8.3e-275; Indels 0; Gaps 0;
Matches 1076; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTGCAAAAACGCAAGAGATTCATTCAGCTGAGGTGTGCGCATCCAGGACGACCTG 60
DB 101 GTGCAAAAACGCAAGAGATTCATTCAGCTGAGGTGTGCGCATCCAGGACGACCTG 160
QY 61 GAACCGATGATTAAGGTATGCGGCTTCGAAATGCGGACGATCTAAATCCCTCTT 120
DB 161 GAACCGATGATTAAGGTATGCGGCTTCGAAATGCGGACGATCTAAATCCCTCTT 220
QY 121 GGTGAGGTTTGGGCGATGACCGATGGATGGATTCCTTTGATTCATCAATAC 180

DB 221 GGTGAGGTTTGGGCGATGACCGATGGATGGATTCCTTTGATTCATCAATAC 280
QY 181 GGCTACGAATGATGGGACGCCCCCATCTGTTTCCGGCTGATTTTGGCGGCTCCACCGAA 240
DB 281 GGCTACGAATGATGGGACGCCCCCATCTGTTTCCGGCTGATTTTGGCGGCTCCACCGAA 340
QY 241 ATGCTGATCATGCGCTCGTGTGGGCGACGCCCCCTGGGCGCATGCGCTTACCA 300
DB 341 ATGCTGATCATGCGCTCGTGTGGGCGACGCCCCCTGGGCGCATGCGCTTACCA 400
QY 301 TTGCTGTAAGTATTCGCGAGATTTCTATGAGTTTCAATTCGCGTGAATGAGTCA 360
DB 401 TTGCTGTAAGTATTCGCGAGATTTCTATGAGTTTCAATTCGCGTGAATGAGTCA 460
QY 361 AACCCCATGCGCGTTCTTATTCGCTTTTCGCGTTATTCGAGCAAGCCTACGAGTCACT 420
DB 461 AACCCCATGCGCGTTCTTATTCGCTTTTCGCGTTATTCGAGCAAGCCTACGAGTCACT 520
QY 421 GCGGCGACGCGCGACGCTGTGCGGCTGGCGCATTTACTCAATGCAATAGCGTTTAC 480
DB 521 GCGGCGACGCGCGACGCTGTGCGGCTGGCGCATTTACTCAATGCAATAGCGTTTAC 580
QY 481 TCCTACTGGGATATTCGCGCGTCTCACCGAGATGCGAGATTCATTCCTTTGAA 540
DB 581 TCCTACTGGGATATTCGCGCGTCTCACCGAGATGCGAGATTCATTCCTTTGAA 640
QY 541 ATTAAGGCGCTGAGTTCGCGCTTGTCTCTCTTTGTGTACGCTGACTTGGATTCCTGC 600
DB 641 ATTAAGGCGCTGAGTTCGCGCTTGTCTCTCTTTGTGTACGCTGACTTGGATTCCTGC 700
QY 601 CGAAGCAAAAAGCAGATCCCTTCTGCTGCTGCGAGTTTGAGCTTCAACATTCGCTT 660
DB 701 CGAAGCAAAAAGCAGATCCCTTCTGCTGCTGCGAGTTTGAGCTTCAACATTCGCTT 760
QY 661 GTGTGATATTCAGATGACGCGCTTATTCGCGGCTGCTGATCTTCTTGATTCCTGC 720
DB 761 GTGTGATATTCAGATGACGCGCTTATTCGCGGCTGCTGATCTTCTTGATTCCTGC 820
QY 721 ATCCGATCTTCTTCTGGAAGAGGCTGTGTAATGACACTGATTTCTCTGATTTCTGC 780
DB 821 ATCCGATCTTCTTCTGGAAGAGGCTGTGTAATGACACTGATTTCTCTGATTTCTGC 880
QY 781 TTGTGTGCGAGTATGAGTCAATTCCTTTGCGCTCGCGCGGCTTCCGTTCTTAATCC 840
DB 881 TTGTGTGCGAGTATGAGTCAATTCCTTTGCGCTCGCGCGGCTTCCGTTCTTAATCC 940
QY 841 TTAAAGCCCTCAGTGATGACAAATTTGTGGGCAAAATGCGAATGAGTCCAGCAGAA 900
DB 941 TTAAAGCCCTCAGTGATGACAAATTTGTGGGCAAAATGCGAATGAGTCCAGCAGAA 1000
QY 901 TCCTTGGCATTTTGAACCGCATCAACGTTTTCGACGAATGCGATGATCTGAAGCTCTAA 960
DB 1001 TCCTTGGCATTTTGAACCGCATCAACGTTTTCGACGAATGCGATGATCTGAAGCTCTAA 1060
QY 961 CCTTTGGTCTCATTTGCGGCTTGGCATTTACAGTGTGGGCGCATCTTTGGCGGTGACGCA 1020
DB 1061 CCTTTGGTCTCATTTGCGGCTTGGCATTTACAGTGTGGGCGCATCTTTGGCGGTGACGCA 1120
QY 1021 CCTTTGAGCGGTGGGCGCTGACGATCTTTTGTGTGACTGGTGAATCTTTTC 1076
DB 1121 CCTTTGAGCGGTGGGCGCTGACGATCTTTTGTGTGACTGGTGAATCTTTTC 1176

RESULT 5
BD014991 1271 bp DNA linear PAT 27-AUG-2002
LOCUS BD014991
DEFINITION Nucleotide sequence encoding the discharge of branched amino acid, method of isolating the same and utilization thereof.
VERSION BD014991.1 GI:22555798
KEYWORDS UP 200118978-A/1.
SOURCE unidentified

ORGANISM unidentified
 unclassified.
 REFERENCE 1 (bases 1 to 1271)
 AUTHORS Kennerknecht, N., Sahm, H., Eggering, L. and Pfeifferle, W.
 TITLE Nucleotide sequence encoding the discharge of branched amino acid,
 method of isolating the same and utilization thereof
 JOURNAL DEUTSCHE HUELIS AG, FORSCHUNGSZENTRUM JUELICH GMBH
 COMMENT OS Corynebacterium glutamicum ATCC14752
 PN JP 2001169788-A/1
 PD 26-JUN-2001
 PR 24-OCT-2000 JP 2000324315
 PI NICOLE KENNERKNECHT, HERMANN SAHM, LOTHAR EGGERING, WALTER PI
 PREFERENCE PC C12N15/09, C07K14/34, C12N1/21, C12P13/06, C12P13/06, C12P13/08//
 PC C12N15/09, C12R1/15, C12N1/21, C12R1/15, C12P13/06, C12R1/15),
 PC C12N15/00,
 PC C12N15/00, C12R1/15)
 CC brnf
 CC brnf
 CC key
 FH gene
 FT gene
 FT gene
 Location/Qualifiers
 1..1271
 /organism="unidentified"
 /mol_type="genomic DNA"
 /db_xref="taxon:32644"

ORIGIN
 Query Match 100.0%; Score 1076; DB 6; Length 1271;
 Best Local Similarity 100.0%; Pred. No. 8.3e-275;
 Matches 1076; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 GTGCAAAAACGCAAGATTTCATTCAGCCCTGAGGCTGCGCCATCCAAAGCAGCCCTG 60
 101 GTGCAAAAACGCAAGATTTCATTCAGCCCTGAGGCTGCGCCATCCAAAGCAGCCCTG 160
 61 GAAACAGATGATTAAGTTATCGGCGCTACGAAATCGCGCAAGGCTTAAAAACCTCCCT 120
 161 GAAACAGATGATTAAGTTATCGGCGCTACGAAATCGCGCAAGGCTTAAAAACCTCCCT 220
 121 GCTGAGGTTTGGGAGTATGATCCGATGATGATGCTGCTCTGTTGATTAATCAATAC 180
 221 GCTGAGGTTTGGGAGTATGATCCGATGATGATGATGCTGCTCTGTTGATTAATCAATAC 280
 181 GGCTCGAATGATGGGAGCCGACGCTGTTCCGGCTGATTTTCGGCGGCTCCACCGAA 240
 281 GGCTCGAATGATGGGAGCCGACGCTGTTCCGGCTGATTTTCGGCGGCTCCACCGAA 340
 241 ATGCTGATGATGCTGCTGTTGAGGCGAGCGCCCTGAGGCGCATGCGCTCACCA 300
 341 ATGCTGATGATGCTGCTGTTGAGGCGAGCGCCCTGAGGCGCATGCGCTCACCA 400
 301 TTGCTGATGATGCTGCTGTTGAGGCGAGCGCCCTGAGGCGCATGCGCTCACCA 360
 401 TTGCTGATGATGCTGCTGTTGAGGCGAGCGCCCTGAGGCGCATGCGCTCACCA 460
 361 AACCCATGCGGTTTCTATTCGTTTTCGGCTTATGACGAAAGCTTACAGCTCACT 420
 461 AACCCATGCGGTTTCTATTCGTTTTCGGCTTATGACGAAAGCTTACAGCTCACT 520
 421 GGGGCGAGGCGCGAGGCTGCTGCGGCGATGATTAATCAATGCAATTAACGTTTAC 480
 521 GGGGCGAGGCGCGAGGCTGCTGCGGCGATGATTAATCAATGCAATTAACGTTTAC 580
 481 TCTTACTGGGATATTCGGCGGCTTCAACCGAGTGGCGATGCGAGGTTGATTCCTTTGA 540
 581 TCTTACTGGGATATTCGGCGGCTTCAACCGAGTGGCGATGCGAGGTTGATTCCTTTGA 640
 541 ATTAAAGGCTGAGTGGCGCTTCTCTCTCTTTGTACAGCTGACTTGGATTCCTGCG 600

Db 641 ATTAAAGGCTGAGTGGCGCTTCTCTCTCTTTGTACAGCTGACTTGGATTCCTGCG 700
 Qy 601 CGAAGCAAAAACGATCCCTTCTCTGCTGCTCGAGGTTTACGTTACCATCTCTCT 660
 Db 701 CGAAGCAAAAACGATCCCTTCTCTGCTGCTCGAGGTTTACGTTACCATCTCTCT 760
 Qy 661 GTGGTATTCAGAGTCAGGCTTATTTGGCGGCTGCTGATCTTCTTGGCTGTGACC 720
 Db 761 GTGGTATTCAGAGTCAGGCTTATTTGGCGGCTGCTGATCTTCTTGGCTGTGACC 820
 Qy 721 ATCCGCTATCTTCTTGGGAAAGGCTGCTTAAATGAACTATTTCTTATCTCC 780
 Db 821 ATCCGCTATCTTCTTGGGAAAGGCTGCTTAAATGAACTATTTCTTATCTCC 880
 Qy 781 TTGTTGCGAGATGTCAGTCATTAATCTTGGCTCGGCGGCTCGGCTCTTAATCC 840
 Db 881 TTGTTGCGAGATGTCAGTCATTAATCTTGGCTCGGCGGCTCGGCTCTTAATCC 940
 Qy 841 TTAAGCCCTACGTGAATCAAAATTTGGGCAAAATGCGATGATGATCCAGAGAA 900
 Db 941 TTAAGCCCTACGTGAATCAAAATTTGGGCAAAATGCGATGATGATCCAGAGAA 1000
 Qy 901 TCTTGCATTTTGAACCGATCAACGTTTCGAGCAATGCGATGATGATGATGATCTAA 960
 Db 1001 TCTTGCATTTTGAACCGATCAACGTTTCGAGCAATGCGATGATGATGATGATCTAA 1060
 Qy 961 CTTTGTGATCATGCGGCTGCGATCAAGTGTGCGGATCTTCTTGGCGGTCAGCGCA 1020
 Db 1061 CTTTGTGATCATGCGGCTGCGATCAAGTGTGCGGATCTTCTTGGCGGTCAGCGCA 1120
 Qy 1021 CTTTGTGATCATGCGGCTGCGATCAAGTGTGCGGATCTTCTTGGCGGTCAGCGCA 1076
 Db 1121 CTTTGTGATCATGCGGCTGCGATCAAGTGTGCGGATCTTCTTGGCGGTCAGCGCA 1176

RESULT 6
 BD014994 1271 bp DNA linear PAT 27-AUG-2002
 LOCUS
 DEFINITION Nucleotide sequence encoding the discharge of branched amino acid,
 method of isolating the same and utilization thereof.
 ACCESSION BD014994.1 GI:22555801
 VERSION JP 2001169788-A/4.
 KEYWORDS
 SOURCE unidentified
 ORGANISM unidentified
 unclassified.
 REFERENCE 1 (bases 1 to 1271)
 AUTHORS Kennerknecht, N., Sahm, H., Eggering, L. and Pfeifferle, W.
 TITLE Nucleotide sequence encoding the discharge of branched amino acid,
 method of isolating the same and utilization thereof
 JOURNAL DEUTSCHE HUELIS AG, FORSCHUNGSZENTRUM JUELICH GMBH
 COMMENT OS Corynebacterium glutamicum ATCC13032
 PN JP 2001169788-A/4
 PD 26-JUN-2001
 PR 24-OCT-2000 JP 2000324315
 PI NICOLE KENNERKNECHT, HERMANN SAHM, LOTHAR EGGERING, WALTER PI
 PREFERENCE PC C12N15/09, C07K14/34, C12N1/21, C12P13/06, C12P13/06, C12P13/08//
 PC C12N15/09, C12R1/15, C12N1/21, C12R1/15, C12P13/06, C12R1/15),
 PC C12N15/00,
 PC C12N15/00, C12R1/15)
 CC brnf
 CC brnf
 CC key
 FH gene
 FT gene
 FT gene
 Location/Qualifiers
 1..1271
 /organism="unidentified"
 /mol_type="genomic DNA"
 /db_xref="taxon:32644"


```

/product="Lrp-like regulator"
/protein_id="AA046687.1"
/db_xref="GI:21311382"
/translation="MKLSDIDRAIIAEISANARISNIALADKVLTPGCLRRVORLE
AEGIIIGASDIHPANRNGREIVDTVLTNFDSESVTAQDEVELHRLRFG
SPDFVRIAGVADLEAYEQLSSHTQTPGIAKISRRAMVVPARQV"

```

ORIGIN

```

Query Match      100.0%; Score 1076; DB 1; Length 2105;
Best local Similarity 100.0%; Pred. No. 8.8e-275;
Matches 1076; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTGCAAAAAGCGAAGATTCATTCAAGCCTGAGGAGGCGCATCGAAGGAGCCCTG 60
DB 1463 GTGCAAAAAGCGAAGATTCATTCAAGCCTGAGGAGGCGCATCGAAGGAGCCCTG 1404
QY 61 GAACCAAGATGATAAGGTTATCGGCGCTACGAAATCGCGCAAGGCTTAAAAACCTCCCT 120
DB 1403 GAACCAAGATGATAAGGTTATCGGCGCTACGAAATCGCGCAAGGCTTAAAAACCTCCCT 1344
QY 121 GCTGCAAGTTGGGCAATACCCGATGATATGCGCTTGGTCTCTGTTATTCATAC 180
DB 1243 GCTGCAAGTTGGGCAATACCCGATGATATGCGCTTGGTCTCTGTTATTCATAC 1284
QY 181 GGCTACGATGATGAGGCGAGCCCACTGTTTCCGCGCTGATTTCCGCGGCTCCACCGAA 240
DB 1283 GGCTACGATGATGAGGCGAGCCCACTGTTTCCGCGCTGATTTCCGCGGCTCCACCGAA 1224
QY 241 ATGCTGATCGCCCTCTGTTGGGCGAGCGCCCTCGGCGCATCGCGCTACACACA 300
DB 1223 ATGCTGATCGCCCTCTGTTGGGCGAGCGCCCTCGGCGCATCGCGCTACACACA 1164
QY 301 TTGCTGATGAACTTCGCGACGATATTCATGCTGTTTCACTCCGCTGATGATGATGATA 360
DB 1163 TTGCTGATGAACTTCGCGACGATATTCATGCTGTTTCACTCCGCTGATGATGATGATA 1104
QY 361 AACCAGATGCGCCCTCTGTTATTCGCTTTCGCGCTTATGACGAAAGCTTACGAGTCACT 420
DB 1103 AACCAGATGCGCCCTCTGTTATTCGCTTTCGCGCTTATGACGAAAGCTTACGAGTCACT 1044
QY 421 GGGGCGAGCCCGCGAGGCTGATGCGCGCGGAGCTTATTCGAAATGACGATGCGCTTAC 480
DB 1043 GGGGCGAGCCCGCGAGGCTGATGCGCGCGGAGCTTATTCGAAATGACGATGCGCTTAC 984
QY 481 TCTACTGGGATATTCGCGGATCTCAACGAGTGGCGATCGAGAATTGATTCCTTTGAA 540
DB 983 TCTACTGGGATATTCGCGGATCTCAACGAGTGGCGATCGAGAATTGATTCCTTTGAA 924
QY 541 ATTAAAGGCGCTCGAGTTGGCCCTTGTCTCTCTTTGCAAGCTGATGATGATGATGATGAT 600
DB 923 ATTAAAGGCGCTCGAGTTGGCCCTTGTCTCTCTTTGCAAGCTGATGATGATGATGATGAT 864
QY 601 CGAAGAAAAGAGATCCCTTCTCTGCTGCTCGAGGTTGAGCTTCAACCATGCTCTT 660
DB 863 CGAAGAAAAGAGATCCCTTCTCTGCTGCTCGAGGTTGAGCTTCAACCATGCTCTT 804
QY 661 GTGGAATATTCAGGCTCAAGCCCTATTTGCGCGCGCTGCTGATCTTCTTGGGCTCTGTAAC 720
DB 803 GTGGAATATTCAGGCTCAAGCCCTATTTGCGCGCGCTGCTGATCTTCTTGGGCTCTGTAAC 744
QY 721 ATCCGACTCTCTTTTGGGAAAAGGCTGCTAAATGACAACTGATTTCTCCGTAATTCCTC 780
DB 743 ATCCGACTCTCTTTTGGGAAAAGGCTGCTAAATGACAACTGATTTCTCCGTAATTCCTC 684
QY 781 TTGTTGTCGAGTATGTCGATCTACTTCTTGGCGCTCCGCGCGGTTCCGTTCTTAATCC 840
DB 683 TTGTTGTCGAGTATGTCGATCTACTTCTTGGCGCTCCGCGCGGTTCCGTTCTTAATCC 624
QY 841 TTAAGCCCTTACGTAATCAATTTGTCGCAAAATGCGCATGTCGATGCGACAGGAA 900
DB 623 TTAAGCCCTTACGTAATCAATTTGTCGCAAAATGCGCATGTCGATGCGACAGGAA 564
QY 901 TCCCTTGCATTTGACCGCATCAACGTTTTCGAGCAATGCGATAGATCTGAAGACTTAA 960

```

```

DB 563 TCCCTTGCATTTGACCGCATCAACGTTTCGAGCAAGCGCATATCTGAAGACTCTAA 504
QY 961 CTTTGGTCTCATTCGCCGTTGGCATATACAGTGTGGCGCATCTTCTTGGCGGTACGCA 1020
DB 503 CTTTGGTCTCATTCGCCGTTGGCATATACAGTGTGGCGCATCTTCTTGGCGGTACGCA 444
QY 1021 CTTTGGTACGCTTGGCGGTGGCGCATATGTTTGTGTTGATGATGATCTTTTC 1076
DB 443 CTTTGTGAGCGGTGGCGGTGGCGCATATGTTTGTGTTGATGATGATCTTTTC 388

RESULT 8
AP005274
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
AUTHORS
TITLE
JOURNAL
COMMENT
FEATURES
source
1. .340000
/organism="Corynebacterium glutamicum ATCC 13032"
/mol_type="genomic DNA"
/strain="ATCC 13032"
/db_xref="taxon:196627"
/note="ATCC 13032"
1. .1575
/gene="Cg10001"
1. .1575
/gene="Cg10001"
/note="PB00308: Bacterial dna protein
TIGR00362: DnaA: chromosomal replication initiator protein
DnaA"
/codon_start=1
/translation="ATPase involved in DNA replication initiation"
/protein_id="BAB97394.1"
/db_xref="GI:21322765"
/translation="MSQNSSSLFTMRQVADLTLSQADSGFNPPTORAYNLT
KPAIDVDAVSTPAAKAVYENDLGLALRVLSLRGRSPSAVSVEPQELPET
PAQGERKYQPDAPVSSNPAKPYEGGGAISTDGMERTSAPEPAPADPE
PELATPQRI.PREPAPNPRVSLNPKYFESFVIGPEPFAAANAVAASPARFN
PLFISGSGSLGTHLHAAGVAGOEPLRIKYVSEFTNDYINSVADDDQEFKR
RKNDDILMWDIOFLAGKEGQEEFPTFNALHQAOKOILSDRPPQLTLEBRL
RRFEGGLITDIOPLERIRIALIMKKQTDGTVDRVLELIASRFESIRELBAL
IRVSAVSLINQIDIKEMALIVLRDILPEPMEITAPVIMETVAEYFESISDTLGA
GKTRAYAHARQAMTICRELITMSLPKIGDVGGKDHITVTMTADKIRQEMERKDY
DETQUTQILIKSRGRN"
/gene="Cg10002"
/complement(1594..1920)
/gene="Cg10002"
/codon_start=1
/translation="Hypothetical protein"

```

```

/protein_id="BAB97395.1"
/db_xref="GI:21322766"
/translation="MIVDNFVYLMIDFKRILSPVCADSPAYPOFLGSCAQGKFRG
ITFPTXKLTGVCNVTAVNKVNDNFVSMAFISGYPQNPCTTPTKLOKPSDGSRAA
ELNPFYI"
gene
2292..3476
/gene="Cg10003"
CDS
2292..3476
/gene="Cg10003"
/EC_number="2.7.7.7"
/note="PF00712: DNA polymerase III beta subunit, N-terminal
domain
PF02767: DNA polymerase III beta subunit, central domain
PF02768: DNA polymerase III beta subunit, C-terminal domain
TIGR00663: dhan: DNA polymerase III, beta subunit"
/codon_start=1
/transl_table=11
/product="DNA polymerase III beta subunit"
/protein_id="BAB97396.1"
/db_xref="GI:21322767"
/translation="MESQVSRVARBEDLVTAAMVARNLPTKPTQVLRAMLITADD
EGLEAGYDYVSTRVLSAEVSGPGRIVAGKLSLETGSLPNKPVDFRIDSKAFV
TCGSRFFELPLIDDPMLPTLPAGTGINAKLFTBAVSQVASAGKDDSLPMLTGV
SMEIVGNCINLAATDFRLALRTPEMEANDPELNKLLIPARTLLDNARSIDGSLNS
IDIAVGTGQVGRGEGFVGTNDRETTRLDADPNVLAIPLLPKHTLAAVEIAPLV
DAIRVSLVAERNQIVAFHSGOVITAGATGAAGAEETPCAFTEGKELTIFANPGY
LKDGSLVPTSRNAVFGTEPRPALIMTEPRPMANENGIFQTPDTYFTLLNPFRL
PG"
gene
3585..4769
/gene="Cg10004"
CDS
3585..4769
/gene="Cg10004"
/note="TIGR00611: recf: Recf protein"
/codon_start=1
/transl_table=11
/product="Recombinational DNA repair APase"
/protein_id="BAB97397.1"
/db_xref="GI:21322768"
/translation="MIRSELDYDRSMPELKVDLEPGITFVIGNGFGKTNIEAIG
YLAHLSHRVSSDAPLRAHAENARSAVANQRELAHLIKPHANQSLNRTKY
RPPRLGIVKTVTLFAPDIALVKGEPARRRYDITATQPRMAGKAYADKTKY
KNALTKATIRARRGTEEGAAALSTLDMDGCLARGAEWAPARLWELBEGKTY
EAVYTIAPSRPAANVKTITDQGSQSEFPAGIIEATLTELAKQRIEISGSL
VGRHDDVLMMDQPAKGFASGETWSPALSLIAENPLKSDTDIILLIDVFSF
LDAGREKLVGIAQVEQVLTIAVHDLPENLKVLTAQHTVTVQDTGTRISILDV
QP"
gene
4766..5302
/gene="Cg10005"
CDS
4766..5302
/gene="Cg10005"
/codon_start=1
/transl_table=11
/product="Hypothetical protein"
/protein_id="BAB97398.1"
/db_xref="GI:21322769"
/translation="MTDPIEQAFERIRAAAMRNSVVDINKNDAFRPPAKGVK
RKKGPSGLDGQKTVRGASLGVLNKEIQRGNGKQIRAGVTSWEEVGAKIA
QHTFVEMIDKCLFITCDSTAWATNLRMWQRIQIVIAEKVPIITELRIFGPAPS
MRKGPLHVGRSPRPDTYG"
gene
5354..7489
/gene="Cg10006"
CDS
5354..7489
/gene="Cg10006"
/EC_number="5.99.1.3"
/note="PF00204: DNA topoisomerase II (N-terminal region)
PF00986: DNA gyrase B subunit, carboxyl terminus
PF01751: Toprim domain
PF02518: Histidine kinase, DNA gyrase B-, phytochrome-like
ATPase
TIGR01055: parE: DNA topoisomerase IV, B subunit
TIGR01058: parE: DNA topoisomerase IV, B subunit
TIGR01059: gyrB: DNA gyrase, B subunit"
/codon_start=1

```

```

/transl_table=11
/product="DNA gyrase (topoisomerase II) B subunit"
/protein_id="BAB97399.1"
/db_xref="GI:21322770"
/translation="MRGTTWGPKRVRKRLYRISSECSLKVANTENHYDASSITILE
GLEAVRPRGQVYSGPRLGHLIMEYDNDVDEMAAGHAKVTVTLLEDGVOYVD
DEKGIYVDHPSGAPVQVYVMTQIAGAGFDDSDYAVSGHGVGISVYNALSTVEA
DKLHKHMYQNEKSPDLELGGNARGTGTTIRFPVPAIFETTFEFTISRLO
EMAFLNKGLITLTDNRADDELEALAEQETLVELSDIDNETVEETTDAPK
KPKRREKKKIFHYPNGLEDYVHYLNKSIHPSIYSEPAKDDHEVEVAQMSNSYK
ESVHTFANTINTERGCTHEGFRSALTSIMNRVAREKLLKEKANLTGDDCREGISA
VISRVGDDPQEGOTKXGNTKIKSFVQVNRNEHGHMLNPAEAKYTIKAVGSA
QRLAARADLRKRSATDILGIPKGLADCSKDEKSELYIVGDSGSGAKGSD
SNQALPLRGLKILNVEKRLDKYLNKAEVQAITTLGLGTHDEEDINKLRHKYLM
ADNDVGQHIATLTLTLFRFMDLVAGHVAIOPPLKXHQKREPOFAYSDEERD
EQLNEGLAAGRKINKDDGIQRYKGLGEMNASLWETMDPTVIRLRVSDIQAORADE
LPSILMGDDVVARSRFTIRNAKDVRLDT"
gene
7830..8798
/gene="Cg10007"
CDS
7830..8798
/gene="Cg10007"
/codon_start=1
/transl_table=11
/product="Hypothetical protein"
/protein_id="BAB97400.1"
/db_xref="GI:21322771"
/translation="MTCYTIKIQSLKNAQREMLKTTLTDTPLGHVTSPEPTTAHV
LPDGSSTPQIVASDNKDSQVLMNPFGRMGCTYTRPLAANKAGFHAIGELAGQ
QSSAKARSKSQWGVHDLASVDFPLQIAAKKALDDEGHPNPLSHSKGQISCLPFA
RPEAEKYNLALFVGAGSGPFRPTFSFKGKGLGAVLLGGIIGHIIVGFPGKVLGK
DLVGYRSGGTNRERMRFRKNSLDLDTAQQIDINVAENKVSIPITFRCRDBDCP
QASIDGLASFVPAQIKMIEIPALGHNWASEPESTVXLPLEQAL"
gene
8795..9466
/gene="Cg10008"
CDS
8795..9466
/gene="Cg10008"
/complement(8795..9466)
/transl_table=11
/product="Hypothetical membrane protein"
/protein_id="BAB97401.1"
/db_xref="GI:21322772"
/translation="MGNPNKKPLDPAQRESLPNRMFVFGIEMFALGITVHAGLG
ITITSSLPVMTASGLSGWTTITNGFNILQCIIVASQRPQVLIILNAPLFGP
LCDSLTQLTWAGTDNVAIVIAVISTILMSIGVTVQVLPNTFAGSGVSAVAK
FNVBERGTMKQIVMTFVVAALISITWGLIGVREGVFAFLFGFVRQWRKLYL
RSIGH"
gene
9471..9914
/gene="Cg10009"
CDS
9471..9914
/gene="Cg10009"
/codon_start=1
/transl_table=11
/product="Hypothetical protein"
/protein_id="BAB97402.1"
/db_xref="GI:21322773"
/translation="MTSLDILVYSTNADAIDMSAFIRETVGKSGISGIALKGAMG
AATKIBSDIYAKSRKLLPEIADSLGILNOBYONGSTADPQHLETPANSSTALDALS
VADKNASINIVPELGVYVGRGKAANKVLEDELPAIGLLEKNAQ"
gene
10104..11177
/gene="Cg10010"
CDS
10104..11177
/gene="Cg10010"
/codon_start=1
/transl_table=11
/product="Hypothetical protein"
/protein_id="BAB97403.1"
/db_xref="GI:21322774"
/translation="MFMIAQRITPIHITAPHLFVARVFIQIRATDARTSLQRLIELS

```

Query Match 100.0%; Score 1076; DB 1; Length 340000;
 Best local similarity 100.0%; Pred. No. 1.5e-274; Indels 0; Gaps 0;
 Matches 1076; Conservative 0; Mismatches 0;

QY 1 GTGCAAAAACGCAAGATTCATTCAAGCCTGAGGTGTGCCATCCAGGACGCCCTG 60
 Db 276829 GTGCAAAAACGCAAGATTCATTCAAGCCTGAGGTGTGCCATCCAGGACGCCCTG 276888
 QY 61 GAACCAAGATGATTAAGTTATCGGCGCTACGAAATCGCGCAAGGCTTAATAAACCCTCCTT 120
 Db 276889 GAACCAAGATGATTAAGTTATCGGCGCTACGAAATCGCGCAAGGCTTAATAAACCCTCCTT 276948
 QY 121 GCTGAGGTTTGGGCGATGATCCGATGTAATGGATTTGGTCTCTGGTTATTCATAC 180
 Db 276949 GCTGAGGTTTGGGCGATGATCCGATGTAATGGATTTGGTCTCTGGTTATTCATAC 277008
 QY 181 GGCTACGAATGTTGGGCGAGCCCACTGTTTCCGGCCGATTTTCGGGGCTCCACCGAA 240
 Db 277009 GGCTACGAATGTTGGGCGAGCCCACTGTTTCCGGCCGATTTTCGGGGCTCCACCGAA 277068
 QY 241 ATGCTGATATCGCCCTGTTGTGGGCGACAGCGCCCTGGGCGCATCGCGCTACGACA 300
 Db 277069 ATGCTGATATCGCCCTGTTGTGGGCGACAGCGCCCTGGGCGCATCGCGCTACGACA 277128
 QY 301 TTGCTGTGAATCTCCGCGAGTATCTATGCGTTTTCATTCGCGCTGCATGTGTCATA 360
 Db 277129 TTGCTGTGAATCTCCGCGAGTATCTATGCGTTTTCATTCGCGCTGCATGTGTCATA 277188
 QY 361 AACCCATTGCGCCGTTCTATTCGGTTTTCGCGCTTATCGACGAAGCTTACGAGTCACT 420
 Db 277189 AACCCATTGCGCCGTTCTATTCGGTTTTCGCGCTTATCGACGAAGCTTACGAGTCACT 277248
 QY 421 GCGGCGAGGCGCGAGCGCTGTGCGGCGTGGCGACTTATCTCAATGCAATAGCGTTTAC 480
 Db 277249 GCGGCGAGGCGCGAGCGCTGTGCGGCGTGGCGACTTATCTCAATGCAATAGCGTTTAC 277308
 QY 481 TCCCTAGGATATTCGGGCGTCTCACCGAGTGGCGATCGAGAGTTGATTCCTTTGAA 540
 Db 277309 TCCCTAGGATATTCGGGCGTCTCACCGAGTGGCGATCGAGAGTTGATTCCTTTGAA 277368
 QY 541 ATTAAGGCGCTGAGTTCGCGCTTGTCTCTTTGTACGCTGACTTGGATTCCTGCG 600
 Db 277369 ATTAAGGCGCTGAGTTCGCGCTTGTCTCTTTGTACGCTGACTTGGATTCCTGCG 277428
 QY 601 CGAAGAAAAGCAATCCCTCTGCTGCTCGAGGTTTGAAGTTACATAGCTT 660
 Db 277429 CGAAGAAAAGCAATCCCTCTGCTGCTCGAGGTTTGAAGTTACATAGCTT 277488
 QY 661 GTGTAATTCAGATCAGAGCTTATTTGCGGCGTGTGATCTTCTGCTGTGACC 720
 Db 277489 GTGTAATTCAGATCAGAGCTTATTTGCGGCGTGTGATCTTCTGCTGTGACC 277548
 QY 721 ATCCGATCTTCTTTTGGGAAAAGCTGTGTAATGACAACTGATTTCTCCGTATTCCTC 780
 Db 277549 ATCCGATCTTCTTTTGGGAAAAGCTGTGTAATGACAACTGATTTCTCCGTATTCCTC 277608
 QY 781 TTGTTGTGAGATGTCAGATCATTTTGTGCGCTCGGCGGTTCCGTTCTTAATCC 840
 Db 277609 TTGTTGTGAGATGTCAGATCATTTTGTGCGCTCGGCGGTTCCGTTCTTAATCC 277668
 QY 841 TTAAAGCCCTTACGTAATCAATTTGTGGGCAAAATGCGCAATGTGATGTCACAGGAA 900
 Db 277669 TTAAAGCCCTTACGTAATCAATTTGTGGGCAAAATGCGCAATGTGATGTCACAGGAA 277728
 QY 901 TCCCTGCAATTTTACGCGATCAAGTTTTCGAGCAATGCAATGATGTAAGACTGTA 960
 Db 277729 TCCCTGCAATTTTACGCGATCAAGTTTTCGAGCAATGCAATGATGTAAGACTGTA 277788
 QY 961 CCTTTGTCTCATTCGCGTGTGATTAAGTGTGCGCATCTTTTGGCGGTGACGCA 1020
 Db 277789 CCTTTGTCTCATTCGCGTGTGATTAAGTGTGCGCATCTTTTGGCGGTGACGCA 277848
 QY 1021 CCTTTGTAGGTTGGCGGTGCGACCATGCTTTTGTGAGTGTGATCTTTTTC 1076
 Db 277849 CCTTTGTAGGTTGGCGGTGCGACCATGCTTTTGTGAGTGTGATCTTTTTC 277904

RESULT 9
 AX120085 349980 bp DNA linear PAT 11-MAY-2001
 LOCUS AX120085
 DEFINITION Sequence 1 from Patent EP1108790.
 ACCESSION AX120085 AX114121
 VERSION AX120085.1 GI:114036800
 KEYWORDS
 SOURCE
 ORGANISM
 Corynebacterium glutamicum
 Corynebacterium glutamicum
 Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 Corynebacterineae; Corynebacteriaceae; Corynebacterium.
 REFERENCE
 1
 AUTHORS Nakagawa, S., Mizoguchi, H., Ando, S., Hayashi, M., Ochiai, K.,
 Yokoi, H., Tateishi, N., Senoh, A., Ikeda, M. and Ozaki, A.
 TITLE Novel polynucleotides
 JOURNAL Patent: EP 1108790-A 1 20-JUN-2001;
 KYOWA HAKKO KOGYO CO., LTD. (JP)
 FEATURES
 source
 1. 349980
 /organism="Corynebacterium glutamicum"
 /mol_type="genomic DNA"
 /db_xref="taxon:1718"
 /note="Seq 1 to long (3,309,400) split in 11, seq 1
 0.000.001 0.349.980 349.980"
 ORIGIN
 Query Match 100.0%; Score 1076; DB 6; Length 349980;
 Best Local Similarity 100.0%; Pred. No. 1,5e-274; Gaps 0;
 Matches 1076; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GTGCAAAAACGCAAGATTCATTCAAGCCTGAGGTGTGCCATCCAGGACGCCCTG 60
 Db 276829 GTGCAAAAACGCAAGATTCATTCAAGCCTGAGGTGTGCCATCCAGGACGCCCTG 276888
 QY 61 GAACCAAGATGATTAAGTTATCGGCGCTACGAAATCGCGCAAGGCTTAATAAACCCTCCTT 120
 Db 276889 GAACCAAGATGATTAAGTTATCGGCGCTACGAAATCGCGCAAGGCTTAATAAACCCTCCTT 276948
 QY 121 GCTGAGGTTTGGGCGATGATCCGATGTAATGGATTTGGTCTCTGGTTATTCATAC 180
 Db 277309 TCCCTAGGATATTCGGGCGTCTCACCGAGTGGCGATCGAGAGTTGATTCCTTTGAA 277368
 QY 181 GGCTACGAATGTTGGGCGAGCCCACTGTTTCCGGCCGATTTTCGGGGCTCCACCGAA 240
 Db 277429 GGCTACGAATGTTGGGCGAGCCCACTGTTTCCGGCCGATTTTCGGGGCTCCACCGAA 277488
 QY 241 ATGCTGATATCGCCCTGTTGTGGGCGACAGCGCCCTGGGCGCATCGCGCTACGACA 300
 Db 277069 ATGCTGATATCGCCCTGTTGTGGGCGACAGCGCCCTGGGCGCATCGCGCTACGACA 277128
 QY 301 TTGCTGTGAATCTCCGCGAGTATCTATGCGTTTTCATTCGCGCTGCATGTGTCATA 360
 Db 277129 TTGCTGTGAATCTCCGCGAGTATCTATGCGTTTTCATTCGCGCTGCATGTGTCATA 277188
 QY 361 AACCCATTGCGCCGTTCTATTCGGTTTTCGCGCTTATCGACGAAGCTTACGAGTCACT 420
 Db 277189 AACCCATTGCGCCGTTCTATTCGGTTTTCGCGCTTATCGACGAAGCTTACGAGTCACT 277248
 QY 421 GCGGCGAGGCGCGAGCGCTGTGCGGCGTGGCGACTTATCTCAATGCAATAGCGTTTAC 480
 Db 277249 GCGGCGAGGCGCGAGCGCTGTGCGGCGTGGCGACTTATCTCAATGCAATAGCGTTTAC 277308
 QY 481 TCCCTAGGATATTCGGGCGTCTCACCGAGTGGCGATGCAATGATGTAAGACTGTA 540
 Db 277309 TCCCTAGGATATTCGGGCGTCTCACCGAGTGGCGATGCAATGATGTAAGACTGTA 277368
 QY 541 ATTAAGGCGCTCAGATTCGCGCTTGTCTCTTTGTGACGCTGACTTGTGAATTCCTG 600
 Db 277369 ATTAAGGCGCTCAGATTCGCGCTTGTCTCTTTGTGACGCTGACTTGTGAATTCCTG 277428
 QY 601 CGAAGAAAAGCAATCCCTCTGCTGCTGCTGCGAGTTTGAAGTTACATAGCTT 660
 Db 277429 CGAAGAAAAGCAATCCCTCTGCTGCTGCTGCGAGTTTGAAGTTACATAGCTT 277488

QY 661 GTGTAATTCAGGTCAGAGCCCTATTGCGGCGTGTGATCTTGGGTCTGTGACC 720
 Db 277489 GTGGTAATTCAGGTCAGAGCCCTATTGCGGCGTGTGATCTTGGGTCTGTGACC 277548
 QY 721 ATCCGATCTTCTTGGGAAAGGCTGTAAATGACAACATGATTTCTCTGATTTCC 780
 Db 277549 ATCCGATCTTCTTGGGAAAGGCTGTAAATGACAACATGATTTCTCTGATTTCC 277608
 QY 781 TTGTTGCGAGATGATGAGTCATTTACTTTGCGCTCCGGGCGGTTCCGTTCTTAATCC 840
 Db 277609 TTGTTGCGAGATGATGAGTCATTTACTTTGCGCTCCGGGCGGTTCCGTTCTTAATCC 277668
 QY 841 TTAACCCCTACGTAAATGACAAATTTGGGCAAAATGGCAATGTGATGTCAGAGGAA 900
 Db 277669 TTAACCCCTACGTAAATGACAAATTTGGGCAAAATGGCAATGTGATGTCAGAGGAA 277728
 QY 901 TCCCTGCAATTTTGAACCGCATGAGTTTCGAGCAATGCGATGATGTAAGACTCTAA 960
 Db 277729 TCCCTGCAATTTTGAACCGCATGAGTTTCGAGCAATGCGATGATGTAAGACTCTAA 277788
 QY 961 CCTTTGCTCATTTGCCGTTGCGATTAAGTGTGGCGCATCTTCTTGGCGGTGAGCA 1020
 Db 277789 CCTTTGCTCATTTGCCGTTGCGATTAAGTGTGGCGCATCTTCTTGGCGGTGAGCA 277848
 QY 1021 CCTTTGAGCGTTGGCGGTGCGACCATGCTTTTGTGTGACTGTGATCTTTTC 1076
 Db 277849 CCTTTGAGCGTTGGCGGTGCGACCATGCTTTTGTGTGACTGTGATCTTTTC 277904

RESULT 10
 AR391954
 LOCUS AR391954 753 bp DNA linear PAT 18-DEC-2003
 DEFINITION Sequence 2 from patent US 6613545.
 ACCESSION AR391954
 VERSION AR391954.1 GI:40115725
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unclassified.
 REFERENCE 1 (bases 1 to 753)
 Kemerhnecht,N., Sahm,H., Eggeling,L. and Pfefferle,W.
 Nucleotide sequences coding for the export of branched amino
 acids, process for the isolation thereof and use thereof
 JOURNAL Patent: US 6613545-A 2 02-SEP-2003;
 FEATURES
 source 1..753
 location/Qualifiers
 /organism="Unknown"
 /mol_type="genomic DNA"

ORIGIN

Query Match 70.0%; Score 753; DB 6; Length 753;
 Best Local Similarity 100.0%; Pred. No. 5.6e-189;
 Matches 753; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTGCAAAAACCGAAGATTCATTCAAGCTGAGGTGTGCCATCCAGGACGCTG 60
 Db 1 GTGCAAAAACCGAAGATTCATTCAAGCTGAGGTGTGCCATCCAGGACGCTG 60
 QY 61 GAACGATGATTAAGGATTCAGGCGCTACGAATTCGCGCAAGGTCTAAAACTCCCTT 120
 Db 61 GAACGATGATTAAGGATTCAGGCGCTACGAATTCGCGCAAGGTCTAAAACTCCCTT 120
 QY 121 GCTGCAAGTTTGGGCGATGTAACCGATGATGATGCTTTGCTCTTGGTTATCAATAC 180
 Db 121 GCTGCAAGTTTGGGCGATGTAACCGATGATGATGCTTTGCTCTTGGTTATCAATAC 180
 QY 181 GGCTACGATGATGAGGCGAGCCCACTGTTTCGCGCTGATTTTGGCGGGCTCCAGCA 240
 Db 181 GGCTACGATGATGAGGCGAGCCCACTGTTTCGCGCTGATTTTGGCGGGCTCCAGCA 240
 QY 241 ATGCTGATGATGCGCTGCTGTTGTGGGCGAGCGCCCTGGGCGCATGCGCTCAACA 300
 Db 241 ATGCTGATGATGCGCTGCTGTTGTGGGCGAGCGCCCTGGGCGCATGCGCTCAACA 300

Db 241 ATGCTGATGATGCGCTGCTGTTGTGGGCGAGCGCCCTGGGCGCATGCGCTCAACA 300
 QY 301 TTGCTGTGAATTCGCGCAGCATTTCTATGAGTTTCAATCCGCTGATGTGTA 360
 Db 301 TTGCTGTGAATTCGCGCAGCATTTCTATGAGTTTCAATCCGCTGATGTGTA 360
 QY 361 AACCCATGCGCGTTTCTATTGCGTTTGGCGCTTATGACGAAGCTTACGAGTCACT 420
 Db 361 AACCCATGCGCGTTTCTATTGCGTTTGGCGCTTATGACGAAGCTTACGAGTCACT 420
 QY 421 GCGGCGAGCGCGCAGGCTGTGCGGCGGCACTTATCTAATGAATAGCGTTTAC 480
 Db 421 GCGGCGAGCGCGCAGGCTGTGCGGCGGCACTTATCTAATGAATAGCGTTTAC 480
 QY 481 TCCCTACTGAGTATTCGCGCGGTCTACCGGAGTGGCATGCGAGAGTTGATTCCTTTGAA 540
 Db 481 TCCCTACTGAGTATTCGCGCGGTCTACCGGAGTGGCATGCGAGAGTTGATTCCTTTGAA 540
 QY 541 ATTAAGGCGCTGAGTTGCGCCCTTGTCTCTCTTGTGACGCTGACTTGGATTCGTC 600
 Db 541 ATTAAGGCGCTGAGTTGCGCCCTTGTCTCTCTTGTGACGCTGACTTGGATTCGTC 600
 QY 601 CGAAGCAAAAAGCAGATTCCTTCTCTGCTGCTGCGAGTTTGAAGCTTCAACATGCTCTT 660
 Db 601 CGAAGCAAAAAGCAGATTCCTTCTCTGCTGCTGCGAGTTTGAAGCTTCAACATGCTCTT 660
 QY 661 GTGTAATTCAGAGTGAAGGCCCTATTGCGGCGCTGATCTTCTTGGGTCTGTGACC 720
 Db 661 GTGTAATTCAGAGTGAAGGCCCTATTGCGGCGCTGATCTTCTTGGGTCTGTGACC 720
 QY 721 ATCCGATCTTCTTGGGAAAGGCTGCTAA 753
 Db 721 ATCCGATCTTCTTGGGAAAGGCTGCTAA 753

RESULT 11
 AX120373
 LOCUS AX120373 753 bp DNA linear PAT 11-MAY-2001
 DEFINITION Sequence 289 from Patent EP1108790.
 ACCESSION AX120373
 VERSION AX120373.1 GI:14037088
 KEYWORDS
 SOURCE Corynebacterium glutamicum
 ORGANISM Corynebacterium glutamicum
 Corynebacterium glutamicum
 Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 Corynebacterinae; Corynebacteriaceae; Corynebacterium.

REFERENCE 1
 Nakagawa,S., Mizoguchi,H., Ando,S., Hayaishi,M., Ochiai,K.,
 Yokoi,H., Tateishi,N., Senoh,A., Ikeda,M. and Ozaki,A.
 Novel polynucleotides
 Patent: EP 1108790-A 289 20-JUN-2001;
 JOURNAL KYOMA HAKKO KOGYO CO., LTD. (JP)
 FEATURES
 source 1..753
 location/Qualifiers
 /organism="Corynebacterium glutamicum"
 /mol_type="unassigned DNA"
 /db_xref="taxon:1718"

ORIGIN

Query Match 70.0%; Score 753; DB 6; Length 753;
 Best Local Similarity 100.0%; Pred. No. 5.6e-189;
 Matches 753; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTGCAAAAACCGAAGATTCATTCAAGCTTGAAGTGTGCGCATCCAGGACGCTG 60
 Db 1 GTGCAAAAACCGAAGATTCATTCAAGCTTGAAGTGTGCGCATCCAGGACGCTG 60
 QY 61 GAACGATGATTAAGGATTCAGGCGCTACGAATTCGCGCAAGGTCTAAAACTCCCTT 120
 Db 61 GAACGATGATTAAGGATTCAGGCGCTACGAATTCGCGCAAGGTCTAAAACTCCCTT 120
 QY 121 GCTGCAAGTTTGGGCGATGTAACCGATGATGATGCTTTGCTCTTGGTTATCAATAC 180
 Db 121 GCTGCAAGTTTGGGCGATGTAACCGATGATGATGCTTTGCTCTTGGTTATCAATAC 180

Db	12	GCTGACAGTTTGGGCATGTACACCGAATGGTATGGGTTTGCTCTTGTTATTCATAAC	180
Qy	181	GGCTACGAATGTTGGGCAGCCCCACACTGTTTTCGGCCCTGATTTTGGCGGCTCCACCGAA	240
Db	181	GGGTACGAATGGTGGGCAGCCCCACACTGTTTTCGGCCCTGATTTTGGCGGCTCCACCGAA	240
Qy	241	ATGCTGGTCATCCGCTCGTGTGGGCGCAGGCGCCCTGGGCGGCATTCGCGCTCACACA	300
Db	241	ATGCTGGTCATCCGCTCGTGTGGGCGCAGGCGCCCTGGGCGGCATTCGCGCTCACACA	300
Qy	301	TTCGTGTGAATTCGCCACGATTTCTATTCGTTTTCAATTCGCGCTGACATGTGTCAA	360
Db	301	TTCGTGTGAATTCGCCACGATTTCTATTCGTTTTCAATTCGCGCTGACATGTGTCAA	360
Qy	361	AACCCCATTTGCCCGTTTCTATTTCGGTTTTCGGCTTATTCAGACGAAGCTTAGCAGTACT	420
Db	361	AACCCCATTTGCCCGTTTCTATTTCGGTTTTCGGCTTATTCAGACGAAGCTTAGCAGTACT	420
Qy	421	GGCGCCAGGCGCCGACGAGTGTGGCGGTGGACATTAATCTCAATGCAAAATGCGTTTCAAC	480
Db	421	GGCGCCAGGCGCCGACGAGTGTGGCGGTGGACATTAATCTCAATGCAAAATGCGTTTCAAC	480
Qy	481	TCTTACTGGTATTCGAGCGGTCTCAACCGAGTGGCGATTCGCAAGTTGATTCCTTTGAA	540
Db	481	TCTTACTGGTATTCGAGCGGTCTCAACCGAGTGGCGATTCGCAAGTTGATTCCTTTGAA	540
Qy	541	ATTAAAGGCGCTCGAATTCGCCCTTTGGCTCTCTTTTGTCAAGCTGACTTTGGAATTCCTGC	600
Db	541	ATTAAAGGCGCTCGAATTCGCCCTTTGGCTCTCTTTTGTCAAGCTGACTTTGGAATTCCTGC	600
Qy	601	CGAACGAAAAGAGATTCCTTCTCTGCTGTCTCGCAGGTTTGACCTTCAACCATTTGCTT	660
Db	601	CGAACGAAAAGAGATTCCTTCTCTGCTGTCTCGCAGGTTTGACCTTCAACCATTTGCTT	660
Qy	661	GTGGTAATTCAGGTCAAGGCCATTTTGGCGGCTGTGATCTTTGGGCTGTGTGACC	720
Db	661	GTGGTAATTCAGGTCAAGGCCATTTTGGCGGCTGTGATCTTTGGGCTGTGTGACC	720
Qy	721	ATCCGGACTTCTTCTTGGGAAAAGCTGTGAAA	753
Db	721	ATCCGGACTTCTTCTTGGGAAAAGCTGTGAAA	753

	RESULT 12				
LOCUS	AX137710				
DEFINITION	Sequence 2 from Patent EP1096010.	753 bp	DNA	linear	PAT 30-MAY-2001
ACCESSION	AX137710				
VERSION	AX137710.1	GI:14273889			
KEYWORDS					
SOURCE					
ORGANISM					
	<i>Corynebacterium glutamicum</i>				
	<i>Corynebacterium glutamicum</i>				
	Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;				
	<i>Corynebacteriaceae</i> ; <i>Corynebacteriaceae</i> ; <i>Corynebacterium</i> .				
REFERENCE					
AUTHORS	Kernerkecht, N., Eggeling, L., Sahm, H. and Pfeifferle, W.				
TITLE	Nucleotide sequences coding for branched-chain amino acids export proteins, method for isolating them and their use				
JOURNAL	Patent: EP 1096010-A 2 02-MAY-2001				
FEATURES					
source	DEGUSSEA AG (DE) ; FORSCHUNGSZENTRUM JUELICH GMBH (DE)				
	Location/Qualifiers				
	1..753				

```
CDS
1..>753
/feature=unannotated protein product; bnf
/codon_start=1
/transl_table=1
/protein_id=CAC39925.1"
/db_xref=GI:14673890"
/db_xref=CA39925:CAC39925"
```

ORIGIN

Query Match	70.0%;	Score 753;	DB 6;	Length 753;
Best Local Similarity	100.0%;	Pred. No. 5.6e-189;		
Matches 753;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

OY	1	GTGCAAAAAAGCAAGAGATCATTTCAAGCTGCAAGGTGCGCATCAAGAGCAAGCCTG	60
Db	1	GTGCAAAAAAGCAAGAGATCATTTCAAGCTGCAAGGTGCGCATCAAGAGCAAGCCTG	60
OY	61	GAACCAAGATGATTAAGGTTATCGGCGCTACGAATCGCGCAAGGCTTAAAACTCCCTT	120
Db	61	GAACCAAGATGATTAAGGTTATCGGCGCTACGAATCGCGCAAGGCTTAAAACTCCCTT	120
OY	121	GCTCAGAGTTGGGCATGTACCCGATTTGGATATGACGTTGGTCTCTGGTTATTAATCAATAC	180
Db	121	GCTCAGAGTTGGGCATGTACCCGATTTGGATATGACGTTGGTCTCTGGTTATTAATCAATAC	180
OY	181	GGTACGAATGTTGGGAGGAGCCCACTGTTTTGCGGCTGATTTTGGCGGCTCCACGAA	240
Db	181	GGTACGAATGTTGGGAGGAGCCCACTGTTTTGCGGCTGATTTTGGCGGCTCCACGAA	240
OY	241	ATGTGTGATATGCGCCCTGTTGTGGGAGGAGCGCCCTGGGCGGATCGGCGTACACA	300
Db	241	ATGTGTGATATGCGCCCTGTTGTGGGAGGAGCGCCCTGGGCGGATCGGCGTACACA	300
OY	301	TTTGTGTGAACTTCCGCCAGATATTTCTATGCGTTTTCATTCGCCGTGCAATGTGTCAA	360
Db	301	TTTGTGTGAACTTCCGCCAGATATTTCTATGCGTTTTCATTCGCCGTGCAATGTGTCAA	360
OY	361	AAACCCATTTGCCGTTTCTATTCGCTTTTGGCGTTATTCAGCAAGGCTACGAGTCACT	420
Db	361	AAACCCATTTGCCGTTTCTATTCGCTTTTGGCGTTATTCAGCAAGGCTACGAGTCACT	420
OY	421	GCGGCCAGGCGCGAGGCTGGTCGGCGCTGGCGCAATTATCTCAATGCAATAGCCTTTAC	480
Db	421	GCGGCCAGGCGCGAGGCTGGTCGGCGCTGGCGCAATTATCTCAATGCAATAGCCTTTAC	480
OY	481	TCTACTGGGGATTTCCGCGGCTTCAACCGAGATGCGCAATCGCAGTGAATTCCTTTGAA	540
Db	481	TCTACTGGGGATTTCCGCGGCTTCAACCGAGATGCGCAATCGCAGTGAATTCCTTTGAA	540
OY	541	ATTAAAGGAGCTCGAGTTGCGCCCTTGTCTCTCTTTTCAAGCTGCACTTTGGATTCCTGC	600
Db	541	ATTAAAGGAGCTCGAGTTGCGCCCTTGTCTCTCTTTTCAAGCTGCACTTTGGATTCCTGC	600
OY	601	CGAACGAAAAAGCAGATCCCTTCTCTGCTGCTCGCAGATTTGAACCTTACCATTTGCTTT	660
Db	601	CGAACGAAAAAGCAGATCCCTTCTCTGCTGCTCGCAGATTTGAACCTTACCATTTGCTTT	660
OY	661	GTGTATTAATCAAGGTCAAGCCCTATTTTGGCGCGCTGCTGATCTTTTGGGTCTGTACCC	720
Db	661	GTGTATTAATCAAGGTCAAGCCCTATTTTGGCGCGCTGCTGATCTTTTGGGTCTGTACCC	720
OY	721	ATCCGGTACTTCTTTCTTGGGAAAAAGCTGCTAAA	753
Db	721	ATCCGGTACTTCTTTCTTGGGAAAAAGCTGCTAAA	753

RESULT 13	
BD014992	
LOCUS	753 bp DNA linear PAT 27-AUG-2002
DEFINITION	Nucleotide sequence encoding the discharge of branched amino acid, method of isolating the same and utilization thereof.
ACCESSION	BD014992
VERSION	BD014992.1 GI:2255799
KEYWORDS	JP 2001169788-A/2.
SOURCE	unidentified

ORGANISM	unclassified
REFERENCE	1 (bases 1 to 753)
AUTHORS	Kennernkuehnto,N., Sahm,H., Eggering,L. and Pfeifferle,W.
TITLE	Nucleotide sequence encoding the discharge of branched amino acid, method of isolating the same and utilization thereof
JOURNAL	Patent: JP 200169788-A 2 26-JUN-2001;
COMMENT	DEUSSA HUELS AG,FORSCHUNGSZENTRUM JUELICH GMBH OS Corynebacterium glutamicum ATCC14752 PN JP 200169788-A/2 PD 26-JUN-2001 PF 24-OCT-2000 JP 2000324315 PR 27-OCT-1999 DE 19951708.8 PI NICOLE KENNERKUEHNHITO,HERRMANN SAHM,LOTHAR EGGERING,WALTER PFEIFFERLE
FEATURES	PC C12N15/09,C07K14/34,C12N1/21,C12P13/06,C12P13/06,C12P13/08// PC (C12N15/09,C12R1:15),(C12N1/21,C12R1:15),(C12P13/06,C12R1:15), PC C12N15/00, PC (C12N15/00,C12R1:15) CC brnp FH Key FT CDS Location/Qualifiers 1..753 Location/Qualifiers /organism="unclassified" /mol_type="genomic DNA" /db_xref="taxon:32644"
ORIGIN	
Query Match	70.0%; Score 753; DB 6; Length 753;
Best Local Similarity	100.0%; Pred. No. 5.6e-169;
Matches	753; Conservative 0; Mismatches 0; Indels 0; Gaps 0
QY	1 GTGCAAAAACCGCAAGATTCATTCAAGCCGTGAGGTGTGCGCATCCAGGCAAGCCCTG 60
DB	1 GTGCAAAAACCGCAAGATTCATTCAAGCCGTGAGGTGTGCGCATCCAGGCAAGCCCTG 60
QY	61 GAACAGATGATAAAGGTTATGGGGCGCTACGAAATCGGCGAAGCTTAAACCTCCCTT 120
DB	61 GAACAGATGATAAAGGTTATGGGGCGCTACGAAATCGGCGAAGCTTAAACCTCCCTT 120
QY	121 GCTGAGGTTTGGGATGATGACCCGATTGGTATGGGTTTGTCTCTTGTTATTCATATC 180
DB	121 GCTGAGGTTTGGGATGATGACCCGATTGGTATGGGTTTGTCTCTTGTTATTCATATC 180
QY	181 GGCTACGAATGTTGGGACAGCCCACTGTTTTCGGCTGATATTTGGCGGGCTCCACCGAA 240
DB	181 GGCTACGAATGTTGGGACAGCCCACTGTTTTCGGCTGATATTTGGCGGGCTCCACCGAA 240
QY	241 ATGCTGATATGAGCCCTGTTGAGGCGAGAGCCCTGGGCGCATCGAGCTCACACA 300
DB	241 ATGCTGATATGAGCCCTGTTGAGGCGAGAGCCCTGGGCGCATCGAGCTCACACA 300
QY	301 TTGCTGTGTAACCTTCGCGACAGTATCTATGCGTTTCAATTCGCCGTGATGTGTCAA 360
DB	301 TTGCTGTGTAACCTTCGCGACAGTATCTATGCGTTTCAATTCGCCGTGATGTGTCAA 360
QY	361 AACCCATGAGCCGTTTCTATTTCGAGTTTTCGGAGCTATGAGCAAGACTACAGACTACT 420
DB	361 AACCCATGAGCCGTTTCTATTTCGAGTTTTCGGAGCTATGAGCAAGACTACAGACTACT 420
QY	421 GCGGCGACGAGCCGCGAGGCTGTGTGCGGTGCGCATTAATCTCAATGCAATAGCGTTTCA 480
DB	421 GCGGCGACGAGCCGCGAGGCTGTGTGCGGTGCGCATTAATCTCAATGCAATAGCGTTTCA 480
QY	481 TCTTACTGAGTATTGGCGAGTCTCACCGAGTGGCATGCAAGATGATATCTTTGAA 540
DB	481 TCTTACTGAGTATTGGCGAGTCTCACCGAGTGGCATGCAAGATGATATCTTTGAA 540
QY	541 ATTAAAGGCTCGAGTTGCGCCTTGTCTCTCTTTGTACGCTGAATTTGATTTCTTG 600
DB	541 ATTAAAGGCTCGAGTTGCGCCTTGTCTCTCTTTGTACGCTGAATTTGATTTCTTG 600

Oy	60	CGAAGCAAAAAGAGATCCCTTCCTCTGCTGCTCGACAGTTTGAGCTTCAACATGTGCTTT	660
Db	601	CGAAGCAAAAAGAGATCCCTTCCTCTGCTGCTCGACAGTTTGAGCTTCAACATGTGCTTT	660
Oy	661	GTGTGTAATTCAGGTCAGGCGCCCTAATTGGCGCGCTGCTGATCTTTCTTGGGTCGTGAC	720
Db	661	GTGTGTAATTCAGGTCAGGCGCCCTAATTGGCGCGCTGCTGATCTTTCTTGGGTCGTGAC	720
Oy	721	ATCCGACTACTTCTCTTGAGAAAGGCTGCTAAA	753
Db	721	ATCCGACTACTTCTCTTGAGAAAGGCTGCTAAA	753
RESULT 14			
BD162490		753 bp	DNA
LOCUS			linear
DEFINITION	BD162490		PAT 17-JAN-2003
ACCESSION	BD162490.1		
VERSION	GI:27868248		
KEYWORDS	JP 2002191370-A/289.		
SOURCE	unidentified		
ORGANISM	unclassified.		
REFERENCE	1 (bases 1 to 753)		
AUTHORS	Nakagawa,S., Mizoguchi,H., Ando,S., Hayashi,M., Ochiai,K., Yokoi,H., Tateishi,N., Senoo,A., Ikeda,M. and Ozaki,A.		
TITLE	Novel polynucleotide		
JOURNAL	Patent: JP 2002191370-A 289 09-JUL-2002;		
COMMENT	KIYOMA HAKKO KOGYO CO LTD		
	OS Corynebacterium glutamicum		
	PN JP 2002191370-A/289		
	PD 09-JUL-2002		
	PF 15-DEC-2000 JP 2000405096		
	PI SATOSHI NAKAGAWA,HIROSHI MIZOGUCHI,SEIKO ANDO,MIKIO HAYASHI, PI KEIICHI OCHIAI		
	PI HARUHIKO YOKOI,NAOKO TATEISHI,AKIHIRO SENOO,MASATO IKEDA,AKIO		
	PI OZAKI		
	PC C12N15/09,C07K14/34,C07K16/12,C07K16/40,C12M1/00, PC C12N1/15,		
	PC C12N1/19,C12N1/21,C12N5/10,C12N9/00,C12N9/02,C12P7/40,C12P13/		
	PC 04,C12P13/08,		
	PC C12P19/00,C12P19/34,C12P21/02,C12Q1/37,C12Q1/68,G01N33/53, PC G01N33/566,		
	PC G01N33/569,G01N33/68,G01N37/00//C12P21/08,(C12N1/21,C12R1/15),		
	PC (C12N1/21,C12R1/13),(C12N1/21,C12R1/01),(C12P13/08,C12R1/15),		
	PC C12N15/00,		
	PC C12N15/00,C12N15/00		
	CC Novel polynucleotide		
	CC Novel polynucleotide		
	FM key		
	Location/Qualifiers		
	FT source 1..753		
FEATURES			
source	Location/Qualifiers		
	1..753		
	/organism="Corynebacterium glutamicum".		
	/organism="unidentified"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:32644"		
ORIGIN			
Query Match	70.0%; Score 753; DB 6; Length 753;		
Best Local Similarity	100.0%; Pred. No. 5,6e-189; Indels 0; Gaps 0		
Matches	753; Conservative 0; Mismatches 0; Indels 0; Gaps 0		
Oy	1	GTGCAAAAACGCAAGAGATTCATTCAAAGCCTGAGGTCGTGCGCATCCAGGACGCTG	60
Db	1	GTGCAAAAACGCAAGAGATTCATTCAAAGCCTGAGGTCGTGCGCATCCAGGACGCTG	60
Oy	61	GAAACAGATGATTAAGGTTATCGGCGCTACGAAATCGGCAAGGCTAAAAACCTCCCT	120
Db	61	GAAACAGATGATTAAGGTTATCGGCGCTACGAAATCGGCAAGGCTAAAAACCTCCCT	120
Oy	121	GCTGCAAGTTGGGAGATGATACCCATGCGTATGTCGCTTGGCTCTTGATTTCAATAC	180
Db	121	GCTGCAAGTTGGGAGATGATACCCATGCGTATGTCGCTTGGCTCTTGATTTCAATAC	180

QY 181 GGCTAACGAATGATGAGGACGCCCCACTGTTTCCGAGCTGATTTTCGGGGGCTCCACCGAA 240
 Db 181 GGCTAACGAATGATGAGGACGCCCCACTGTTTCCGAGCTGATTTTCGGGGGCTCCACCGAA 240
 QY 241 ATGCTGATCATGCGCCCTGTTGTGGGCGCAGCGCCCGGCGCATATGCGCTCACACA 300
 Db 241 ATGCTGATCATGCGCCCTGTTGTGGGCGCAGCGCCCGGCGCATATGCGCTCACACA 300
 QY 301 TTGCTGTGAATCTTCGCGCAGATTCATGATGCTTTTCATTCGCGCTGATGTCGAAA 360
 Db 301 TTGCTGTGAATCTTCGCGCAGATTCATGATGCTTTTCATTCGCGCTGATGTCGAAA 360
 QY 361 AACCCATTTGCCCGCTTCTATTCGCTTTTCGCGCTTATGCGAAGGCTATGCGATGCT 420
 Db 361 AACCCATTTGCCCGCTTCTATTCGCTTTTCGCGCTTATGCGAAGGCTATGCGATGCT 420
 QY 421 GCGGCGAGGCGCGCAGGCTGCTGCGCGTGGCGACTTATCTCAATGCAATAGCGTTTAC 480
 Db 421 GCGGCGAGGCGCGCAGGCTGCTGCGCGTGGCGACTTATCTCAATGCAATAGCGTTTAC 480
 QY 481 TCCCTACTGGATATTTGGGCGGTCTGACCGGAGTGGGAGATCGGAGAGTTGATTCCTTTGAA 540
 Db 481 TCCCTACTGGATATTTGGGCGGTCTGACCGGAGTGGGAGATCGGAGAGTTGATTCCTTTGAA 540
 QY 541 ATTAAGGCGCTGAGTTCGCGCTTCTCTCTCTTTTGCACGCTGACTTGATTCCTGCG 600
 Db 541 ATTAAGGCGCTGAGTTCGCGCTTCTCTCTCTTTTGCACGCTGACTTGATTCCTGCG 600
 QY 601 CGAAGGAAAGGCAATCCCTTCTCTGCTGCTGCGAGGTTGAGCTTACCATTTGCTCTT 660
 Db 601 CGAAGGAAAGGCAATCCCTTCTCTGCTGCTGCGAGGTTGAGCTTACCATTTGCTCTT 660
 QY 661 GTGATATTCAGGTCAGGCGCTATTTTGGGCGCTGATCTTCTTGGTCTGTGAGC 720
 Db 661 GTGATATTCAGGTCAGGCGCTATTTTGGGCGCTGATCTTCTTGGTCTGTGAGC 720
 QY 721 ATCCGTAATCTTCTTCTTGGGAAAGGCTGCTAAA 753
 Db 721 ATCCGTAATCTTCTTCTTGGGAAAGGCTGCTAAA 753

RESULT 15

AR391955

LOCUS AR391955 324 bp DNA linear PAT 18-DEC-2003
 DEFINITION Sequence 4 from patent US 6613545.
 ACCESSION AR391955
 VERSION AR391955.1 GI:40115726
 KEYWORDS

SOURCE Unknown.

ORGANISM Unknown.

REFERENCE Unclassified.

1 (bases 1 to 324)

Kemerzinech,N., Salm,H., Eggeling,L. and Pfeifferle,W.

Nucleotide sequences coding for the export of branched chain amino

acids, process for the isolation thereof and use thereof

Patent: US 6613545-A 4 02-SEP-2003;

Location/Qualifiers

1..324

/organism="unknown"

/mol_type="genomic DNA"

ORIGIN

Query Match 30.1%; Score 324; DB 6; Length 324;

Best Local Similarity 100.0%; Pred. No. 5.5e-75;

Matches 324; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 753 ATGACAACTGATTTCTCTGATTTCTCTGTTGTGCGAGTATGCAATGATTAATCTTT 812

Db 1 ATGACAACTGATTTCTCTGATTTCTCTGTTGTGCGAGTATGCAATGATTAATCTTT 60

QY 813 GGGCTCCGGGCGTTCCTTATCTTAAGCCCTAAGCAATGATCAATTTGTGGC 872

Db 61 GCGCTCCGGGCGTTCCTTATCTTAAGCCCTAAGCAATGATTAATGAGGC 120
 QY 873 AAAATGCCATGTGATGCGCAGCGAGATCTTGCCATTTTACCGCATCAAGTTTGGC 932
 Db 121 AAAATGCCATGTGATGCGCAGCGAGATCTTGCCATTTTACCGCATCAAGTTTGGC 180
 QY 933 AGCAATGCCATGTGATGCGCAGCGAGATCTTGCCATTTTACCGCATCAAGTTTGGC 992
 Db 181 AGCAATGCCATGTGATGCGCAGCGAGATCTTGCCATTTTACCGCATCAAGTTTGGC 240
 QY 993 GTGGCGCATCTTCTTGGCGGTGCGACGCACTTGTGAGCGTTGGCGTGGCAATGCTT 1052
 Db 241 GTGGCGCATCTTCTTGGCGGTGCGACGCACTTGTGAGCGTTGGCGTGGCAATGCTT 300
 QY 1053 TTGTTGCACTGCTGAATCTTTTC 1076
 Db 301 TTGTTGCACTGCTGAATCTTTTC 324

Search completed: April 19, 2004, 18:29:53
 Job time: 4563.51 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 19, 2004, 10:56:53 ; Search time 501.089 Seconds

(without alignments)
9122.247 Million cell updates/sec

Title: US-10-608-504-1_COPY_101_1176

Sequence: 1 gtgcacaaacgcgaagatc.....ctgacgctgacatttc 1076

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

N_Geneseq_29Jan04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001s:*
5: geneseqn2002s:*
6: geneseqn2003s:*
7: geneseqn2003bs:*
8: geneseqn2003cs:*
9: geneseqn2003cs:*
10: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1076	100.0	1271	4	AAH21109 C. glutam
2	1076	100.0	1271	4	AAH21112 C. glutam
3	1076	100.0	349980	5	AAH64966 C. glutam
4	753	70.0	753	4	AAH21110 C. glutam
5	753	70.0	753	5	AAH65254 C. glutam
6	553	51.4	564	7	ACA01969 C. glutam
7	324	30.1	324	4	AAH21111 C. glutam
8	324	30.1	324	5	AAH65255 C. glutam
9	324	30.1	324	7	ACA01970 C. glutam
10	273.6	25.4	397	4	AAH61693 C. glutam
11	273.6	25.4	778	4	AAH61695 C. glutam
12	127.4	11.8	711	5	AAH68373 C. glutam
13	127.4	11.8	711	7	ACA01294 C. glutam
14	127.4	11.8	309400	5	AAH68534 C. glutam
15	106	9.9	177	7	ACA01968 C. glutam
16	85	7.9	715	7	AAH61688 C. glutam
17	48	4.5	10732	3	AAH10594 Gene enco
18	47	4.4	2000	7	ADA71938 Rice gene
19	40.8	3.8	7645	9	ADBS9610 Human gen
20	40.2	3.7	4947	2	AAH69627 Tspst17 po
21	39.8	3.7	768	3	ADG91639 E. faeciu
22	39.2	3.6	6741	3	AAH10595 Gene enco
23	37.6	3.5	759	7	ACF69125 Photornab

C	24	37.6	3.5	5059	2	AAH64332
	25	37.6	3.5	110000	7	ACF67367_20
	26	37.6	3.5	243072	7	ACF65382
	27	37.2	3.5	1024	6	ABX66366
	28	36.8	3.4	1230	7	ACA42205
	29	36.4	3.4	751	2	AAH68489
	30	36.4	3.4	751	2	AAH61146
	31	36.4	3.4	751	3	AAH66252
	32	36.4	3.4	751	3	ABH71157
	33	36.4	3.4	751	4	AAH93368
	34	36.4	3.4	751	4	AAH63460
	35	36.4	3.4	751	4	AAH64682
	36	36.4	3.4	751	5	ACA59269
	37	36.4	3.4	751	5	AAH10011
	38	36.4	3.4	751	6	ABH94832
	39	36.4	3.4	751	6	ABH58541
	40	36.4	3.4	751	7	ACG94996
	41	36.4	3.4	751	9	ADH13462
	42	36.4	3.3	4590	5	ADA71938
	43	35.6	3.3	4590	5	AAH24065
	44	35.6	3.3	135638	7	ABX34289
	45	35.6	3.3	135638	7	ABX34289

ALIGNMENTS

RESULT 1
ID AAH21109
AAH21109 standard; DNA; 1271 BP.
XX
AC AAH21109;
XX
DT 05-SEP-2001 (first entry)
XX
DE C. glutamicum DNA encoding brnF and brnE.
XX
KW L-amino acid production; brnF; brnE; branched-chain amino acid;
KW coryneform bacterium; leucine; isoleucine; valine; medicine;
KW animal nutrition; ds.
XX
OS Corynebacterium glutamicum.
XX
FH Key Location/Qualifiers
FT CDS 101..856
FT FT /*tag= a
FT FT /product= "brnF"
FT CDS 853..1179
FT FT /*tag= b
FT FT /product= "brnE"
XX
XX EPI096010-A1.
XX
XX 02-MAY-2001.
XX
XX 11-OCT-2000; 2000EP-00122057.
XX
XX 27-OCT-1999; 99DE-01051708.
XX
XX (DEGS) DEGUSSA AG.
XX (KERJ) FORSCHUNGSZENTRUM JUELICH GMBH.
XX
XX Kennernrecht N, Eggeling L, Sahm H, Pfefflerle W;
XX WPI; 2001-391595/42.
XX P-PSDB; AAB86247; AAB86248.
XX
XX New export genes from coryneform bacteria, useful for increasing
XX fermentative production of branched-chain amino acids.
XX
XX Claim 4 (i); Page 13; 23pp; German.
XX
XX This invention describes a novel isolated polynucleotide (i) containing

CC at least one sequence that (i) is 70% identical with a sequence that
 CC encodes at least one of 251 amino acid (aa) (3) or 108 aa (5)
 CC polypeptides; (ii) encodes a polypeptide at least 70% identical with (3)
 CC or (5); (iii) is the complement of (i) or (ii), or (iv) contains at least
 CC 15 consecutive bases from (1)-(11). The invention also describes (a)
 CC protein derived from sequences of 753 bp (2) or 324 bp (4); (b)
 CC coryneform microorganisms, especially Corynebacterium, transformed with
 CC one or more (i), where these are replicative DNA; (c) production of
 CC branched-chain L-aa by fermentation of coryneform bacteria in which the
 CC bnf and/or bnf genes (or equivalent sequences) are amplified,
 CC especially overexpressed; and (d) method for isolating the bnf and/or
 CC bnf genes. (1) is used for transformation of coryneform bacteria being
 CC used for fermentative production of branched-chain amino acids,
 CC specifically leucine, isoleucine and valine, which are useful in medicine
 CC and animal nutrition. (i) can also be used as source of primers and
 CC probes for isolation of related sequences. Transformation with (i)
 CC increases yield of branched-chain amino acids. This sequence encodes the
 CC corynebacterium glutamicum ATCC 14752 bnf and bnf proteins described in
 CC the method of the invention

XX SQ Sequence 1271 BP; 265 A; 328 C; 318 G; 360 T; 0 U; 0 Other;

Query Match 100.0%; Score 1076; DB 4; Length 1271;
 Best Local Similarity 100.0%; Pred. No. 0; Mismatches 0; Gaps 0;
 Matches 1076; Conservative 0; Indels 0;

QY 1 GTGCAAAAACGACAGATTCATTCAGGCTGGAGGTGCGCATCCAGGACGCTG 60
 Db 101 GTGCAAAAACGACAGATTCATTCAGGCTGGAGGTGCGCATCCAGGACGCTG 160
 QY 61 GAACCAAGATTAAGTTATCGGCGCTACGAAATGCGCAAGTATAAAGCTCCCT 120
 Db 161 GAACCAAGATTAAGTTATCGGCGCTACGAAATGCGCAAGTATAAAGCTCCCT 220
 QY 121 GGTGAGGTTGGGCAATGTAACCGATTTGATGCTGTTGCTGTTATTCATAC 180
 Db 221 GGTGAGGTTGGGCAATGTAACCGATTTGATGCTGTTGCTGTTATTCATAC 280
 QY 181 GGTGAGGTTGGGCAATGTAACCGATTTGATGCTGTTGCTGTTATTCATAC 240
 Db 281 GGTGAGGTTGGGCAATGTAACCGATTTGATGCTGTTGCTGTTATTCATAC 340
 QY 241 ATGCTGATCGGCTGCTGTTGTTGGGCGACGCGCTGAGGCTGCTGCTACCA 300
 Db 341 ATGCTGATCGGCTGCTGTTGTTGGGCGACGCGCTGAGGCTGCTGCTACCA 400
 QY 301 TTGCTGATGAACTTCGCGCAAGTATTCATGCTGTTGCTGCTGCTGCTGCT 360
 Db 401 TTGCTGATGAACTTCGCGCAAGTATTCATGCTGTTGCTGCTGCTGCTGCT 460
 QY 361 AACCCGATGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
 Db 461 AACCCGATGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 520
 QY 421 GCGGCGAGGCGCGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
 Db 521 GCGGCGAGGCGCGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 580
 QY 481 TCCTACTGGGATTCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
 Db 581 TCCTACTGGGATTCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 640
 QY 541 ATTAAAGGCTCGAGTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
 Db 641 ATTAAAGGCTCGAGTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 700
 QY 601 CGAAGCAAAAAGAGATCCCTTCTGCGGCTGCGAGGTTGAGCTTCAACATGCTCT 660
 Db 701 CGAAGCAAAAAGAGATCCCTTCTGCGGCTGCGAGGTTGAGCTTCAACATGCTCT 760
 QY 661 GTGGATATTCAGAGTCCAGGCTTATTTGCGGCTGCTGCTGCTGCTGCTGCTGCT 720
 Db 761 GTGGATATTCAGAGTCCAGGCTTATTTGCGGCTGCTGCTGCTGCTGCTGCTGCT 820

QY 721 ATCCGGATCTTCTTGGGAAAGGCTGCTAATGCAACTGATTTCTCCTGATTTCTCC 780
 Db 821 ATCCGGATCTTCTTGGGAAAGGCTGCTAATGCAACTGATTTCTCCTGATTTCTCC 880
 QY 781 TTGTTGCGAGATGTCAGTATTAATCTTTGCGCTCGGCGGCTTCGTTCTTAATCC 840
 Db 881 TTGTTGCGAGATGTCAGTATTAATCTTTGCGCTCGGCGGCTTCGTTCTTAATCC 940
 QY 841 TTAGCCCTTACGATTCACATTTTGGGCAAAATGGGAGTGGATGCCAGCGGAA 900
 Db 941 TTAGCCCTTACGATTCACATTTTGGGCAAAATGGGAGTGGATGCCAGCGGAA 1000
 QY 901 TCCTTGCATTTTGGACCGCATCAACGTTTGGCAGCATGGCATGATCTGAAGACTTAA 960
 Db 1001 TCCTTGCATTTTGGACCGCATCAACGTTTGGCAGCATGGCATGATCTGAAGACTTAA 1060
 QY 961 CCTTGGTCTCATTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
 Db 1061 CCTTGGTCTCATTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1120
 QY 1021 CCTTGGTGGAGGCTGGGCTGCGACATCGTTTGTGGAGTGGATCTTTTC 1076
 Db 1121 CCTTGGTGGAGGCTGGGCTGCGACATCGTTTGTGGAGTGGATCTTTTC 1176

RESULT 2
 AAH21112
 ID AAH21112 standard; DNA; 1271 BP.
 XX AC
 XX AAH21112;
 DT 05-SEP-2001 (first entry)
 XX C. glutamicum DNA encoding bnf and bnf.
 DE L-amino acid production; bnf; bnf; branched-chain amino acid;
 XX coryneform bacterium; leucine; isoleucine; valine; medicine;
 KM animal nutrition; ds.
 XX Corynebacterium glutamicum.
 OS
 FH Key Location/Qualifiers
 FT CDS 101..856
 FT /tag= a
 FT /product= "bnf"
 FT 853..1179
 FT /tag= b
 FT /product= "bnf"

EP1096010-A1.
 PD 02-MAY-2001.
 XX 11-OCT-2000; 2000EP-00122057.
 PF 27-OCT-1999; 99DE-01051708.
 XX (DEGS) DEGUSSA AG.
 PA (KERJ) FORSCHUNGSZENTRUM JUELICH GMBH.
 XX Kennerknecht N, Eggeling L, Sahn H, Pfeifferle W;
 PI WPI; 2001-391595/42.
 DR New export genes from coryneform bacteria, useful for increasing
 XX fermentative production of branched-chain amino acids.
 PT Claim 4 (1); Page 17-18; 23pp; German.
 PS
 XX This invention describes a novel isolated polynucleotide (1) containing
 CC at least one sequence that (i) is 70% identical with a sequence that
 CC encodes at least one of 251 amino acid (aa) (3) or 108 aa (5)

CC polypeptides; (ii) encodes a polypeptide at least 708 identical with (3)
 CC or (5); (iii) is the complement of (1) or (1i), or (iv) contains at least
 CC 15 consecutive bases from (1)-(iii). The invention also describes (a)
 CC protein derived from sequences of 753 bp (2) or 324 bp (4); (b)
 CC coryneform microorganisms, especially Corynebacterium, transformed with
 CC one or more (1), where these are replicative DNA; (c) production of
 CC branched-chain L-aa by fermentation of coryneform bacteria in which the
 CC bms and/or bms genes (or equivalent sequences) are amplified,
 CC especially overexpressed; and (d) method for isolating the bms and/or
 CC bms genes. (1) is used for transformation of coryneform bacteria being
 CC used for fermentative production of branched-chain amino acids,
 CC specifically leucine, isoleucine and valine, which are useful in medicine
 CC and animal nutrition. (1) can also be used as source of primers and
 CC probes for isolation of related sequences. Transformation with (1)
 CC increases yield of branched-chain amino acids. This sequence encodes the
 CC Corynebacterium glutamicum ATCC 13032 bms and bms proteins described in
 CC the method of the invention

Sequence 1271 BP; 265 A; 328 C; 318 G; 360 T; 0 U; 0 Other;

Query Match 100.0%; Score 1076; DB 4; Length 1271;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1076; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTGCAAAAACGCAAGATTCATTCAAGCTGAGAGGTGTGCGCATCCAAAGCAGCCCTG 60
 DB 101 GTGCAAAAACGCAAGATTCATTCAAGCTGAGAGGTGTGCGCATCCAAAGCAGCCCTG 160
 QY 61 GAACCAATGATTAAGTTATCGGCGCTACGAATTCGCGCAGGTCTTAAACCTCCCTT 120
 DB 161 GAACCAATGATTAAGTTATCGGCGCTACGAATTCGCGCAGGTCTTAAACCTCCCTT 220
 QY 121 GCTGAGGTTGGGATGATCCGATTTGATGCGTTGCTCTTGGTTATTCATATC 180
 DB 221 GCTGAGGTTGGGATGATCCGATTTGATGCGTTGCTCTTGGTTATTCATATC 280
 QY 181 GGTACGATGATGAGGAGCCCACTGTTTCGCGCTGATTTTCGCGGCTCCAGCGAA 240
 DB 281 GGTACGATGATGAGGAGCCCACTGTTTCGCGCTGATTTTCGCGGCTCCAGCGAA 340
 QY 241 ATGCGTGCATCGCCCTGTGTGAGGAGCGCGCCCTGCGCGCATCGCGCTACACCA 300
 DB 341 ATGCGTGCATCGCCCTGTGTGAGGAGCGCGCCCTGCGCGCATCGCGCTACACCA 400
 QY 301 TTGCTGATGAATTCGCGCAGATTTCTATGCGTTTCAATCCCGCTGATGATGTCAAA 360
 DB 401 TTGCTGATGAATTCGCGCAGATTTCTATGCGTTTCAATCCCGCTGATGATGTCAAA 460
 QY 361 AACCCATGCGCCGTTTATTCGCTTTCGCGCTTATCGACGAAAGCTTACGAGTCACT 420
 DB 461 AACCCATGCGCCGTTTATTCGCTTTCGCGCTTATCGACGAAAGCTTACGAGTCACT 520
 QY 421 GCGGCAAGCCCGCAGCTGTGCGCGTGCAGACTTATCTCAATGCAAAATAGCCTTTCAC 480
 DB 521 GCGGCAAGCCCGCAGCTGTGCGCGTGCAGACTTATCTCAATGCAAAATAGCCTTTCAC 580
 QY 481 TCTCTACTGAGTATTCGCGGCTCTCAACCGAGTGCAGATCGACAGTTGATTCCTTTGAA 540
 DB 581 TCTCTACTGAGTATTCGCGGCTCTCAACCGAGTGCAGATCGACAGTTGATTCCTTTGAA 640
 QY 541 ATTAAGGCTCGAGTTGCGCCCTTGTCTCTCTTTTGCAGCGTGAATTTGATTCCTGCG 600
 DB 641 ATTAAGGCTCGAGTTGCGCCCTTGTCTCTCTTTTGCAGCGTGAATTTGATTCCTGCG 700
 QY 601 CGAAGCAAAAACGAGATCTCTCTCTGCTCGAGGTTTGAAGCTTCAACCATTCCTCTT 660
 DB 701 CGAAGCAAAAACGAGATCTCTCTCTGCTCGAGGTTTGAAGCTTCAACCATTCCTCTT 760
 QY 661 GTGGATATTCAGAGTCAAGCCCTATTTTGGCGGCTGATCTTCTTGGGCTGTGAC 720
 DB 761 GTGGATATTCAGAGTCAAGCCCTATTTTGGCGGCTGATCTTCTTGGGCTGTGAC 820
 QY 721 ATCCGTAATCTTCTTCTTGGGAAAGGCTGCTAAATGACAATGATTTCTCTGATTCCTC 780

DB 821 ATCCGTAATCTTCTTCTTGGGAAAGGCTGCTAAATGACAATGATTTCTCTGATTCCTC 880
 QY 781 TTGTTGTGCAATGATGAGTCAATTCATTATTCGCTCCGCGGCTTCGTTTAAATCC 840
 DB 881 TTGTTGTGCAATGATGAGTCAATTCATTATTCGCTCCGCGGCTTCGTTTAAATCC 940
 QY 841 TTAAAGCCCTTACGATATCAATTTTGGGCAAAATGCGATGTGATGCCAGAGAA 900
 DB 941 TTAAAGCCCTTACGATATCAATTTTGGGCAAAATGCGATGTGATGCCAGAGAA 1000
 QY 901 TCTTGTGCAATTTGACCGCATCAAGTTTTCGAGCAATGCGATGATGAGTCTTAA 960
 DB 1001 TCTTGTGCAATTTGACCGCATCAAGTTTTCGAGCAATGCGATGATGAGTCTTAA 1060
 QY 961 CTTTGTGCTCATGCGCTTGCATTAAGTGTGCGCATCTTTTGGCGGCTGACGCA 1020
 DB 1061 CTTTGTGCTCATGCGCTTGCATTAAGTGTGCGCATCTTTTGGCGGCTGACGCA 1120
 QY 1021 CTTTGTGCAATTTGAGCGCTGCGCATCAATGTTTGTGAGTGTGATCTTTC 1076
 DB 1121 CTTTGTGCAATTTGAGCGCTGCGCATCAATGTTTGTGAGTGTGATCTTTC 1176

RESULT 3
 AAH64966
 ID AAH64966 standard; DNA; 349980 BP.

AC AAH64966;
 XX 26-SEP-2001 (first entry)
 DT XX
 DE C glutamicum coding sequence fragment SEQ ID NO: 1.
 KM Coryneform bacterium; amino acid synthesis; vitamin; saccharide;
 XX organic acid synthesis; ds.
 OS Corynebacterium glutamicum.
 XX
 PN EP1108790-A2.
 PD 20-JUN-2001.
 XX
 PF 18-DEC-2000; 2000EP-00127688.
 XX
 PR 16-DEC-1999; 99JP-00377484.
 PR 07-APR-2000; 2000JP-00159162.
 PR 03-AUG-2000; 2000JP-00280988.
 XX
 PA (KYO) KYOWA HAKKO KOGYO KK.
 XX
 PI Nakagawa S, Mizoguchi H, Ando S, Hayaishi M, Ochiai K, Yokoi H;
 PI Tateishi N, Senoh A, Ikeda M, Ozaki A;
 XX
 DR WPI; 2001-376931/40.
 XX
 PT Novel polynucleotides derived from Coryneform bacteria, for identifying
 PT mutation point of a gene, measuring expression of a gene, analyzing
 PT expression profile or pattern of a gene and identifying homologous gene.
 XX
 PS Claim 7; SEQ ID NO 1; 246bp + Sequence Listing; English.
 XX
 CC The present invention provides a number of nucleotide and protein
 CC sequences from the Coryneform bacterium Corynebacterium glutamicum. These
 CC are useful for identifying the mutation point of a gene derived from a
 CC mutant of coryneform bacterium, measuring expression amount and analysing
 CC the expression profile or expression pattern of a gene derived from
 CC Coryneform bacterium, and identifying a homologue of a gene derived from
 CC coryneform bacterium. Coryneform bacteria are useful for producing amino
 CC acids, nucleic acids, vitamins, saccharides and organic acids,
 CC particularly L-lysine. The present sequence is a nucleic acid described
 CC in the exemplification of the invention. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained

CC in electronic format directly from the European Patent Office
 XX Sequence 349980 BP; 79703 A; 91547 C; 98381 G; 80349 T; 0 U; 0 Other;
 SQ Query Match 100.0%; Score 1076; DB 5; Length 349980;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1076; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTGCAAAAACGCAAGATTCATTCAGCCTGAGAGGTGTCGATCCAGAGCCGCTG 60
 DB 276829 GTGCAAAAACGCAAGATTCATTCAGCCTGAGAGGTGTCGATCCAGAGCCGCTG 276888
 QY 61 GAACAGATGATTAAGTTATGCGCGCTACGAAATCGCGCAAGCTTAAACCTCCCTT 120
 DB 276889 GAACAGATGATTAAGTTATGCGCGCTACGAAATCGCGCAAGCTTAAACCTCCCTT 276948
 QY 121 GCTGCAAGTTGGGCGATGATCCAGTGGATGCGTTGGCTCTCTTGGTTATTCATAC 180
 DB 276949 GCTGCAAGTTGGGCGATGATCCAGTGGATGCGTTGGCTCTCTTGGTTATTCATAC 277008
 QY 181 GACTACGATGGTGGGCGAGCCCACTGTTTTCCGGCTGATTTTCGCGGGCTCCAGCGAA 240
 DB 277009 GACTACGATGGTGGGCGAGCCCACTGTTTTCCGGCTGATTTTCGCGGGCTCCAGCGAA 277068
 QY 241 ATGCTGTCATCGCCCTGTTGTGGGCGAGCGCCCTTGGGCGCCATCGCGCTACCA 300
 DB 277069 ATGCTGTCATCGCCCTGTTGTGGGCGAGCGCCCTTGGGCGCCATCGCGCTACCA 277128
 QY 301 TTGCTGTAATTCGCGCAAGTATTCATGCGTTTTCATTCGCGCTGATGCGTCA 360
 DB 277129 TTGCTGTAATTCGCGCAAGTATTCATGCGTTTTCATTCGCGCTGATGCGTCA 277188
 QY 361 AACCCATTTGCCGTTTCTATTCGCTTTCGCGCTTATTCAGAGAGCTTACGAGTCA 420
 DB 277189 AACCCATTTGCCGTTTCTATTCGCTTTCGCGCTTATTCAGAGAGCTTACGAGTCA 277248
 QY 421 GCGGCGAGCGCGGCGGCTGCGCGCTGCGCGCTTATTCAGAGAGCTTACGAGTCA 480
 DB 277249 GCGGCGAGCGCGGCGGCTGCGCGCTGCGCGCTTATTCAGAGAGCTTACGAGTCA 277308
 QY 481 TCCTACTGGGATTCGCGCGCTTCAACCGAGTGGCGATTCGAGTTCATTCCTTTGAA 540
 DB 277309 TCCTACTGGGATTCGCGCGCTTCAACCGAGTGGCGATTCGAGTTCATTCCTTTGAA 277368
 QY 541 ATTAAGGCGCTTGCAGTTCGCTTTCCTTCTTTCGAGCTTACCTTTCGATTCCTG 600
 DB 277369 ATTAAGGCGCTTGCAGTTCGCTTTCCTTCTTTCGAGCTTACCTTTCGATTCCTG 277428
 QY 601 CGAAGCAAAAACGATCCCTTCTGCTGCTCGAGGTTGAGCTTACCATTCCTCTT 660
 DB 277429 CGAAGCAAAAACGATCCCTTCTGCTGCTCGAGGTTGAGCTTACCATTCCTCTT 277488
 QY 661 GTGGTAATTCAGGTCAGGCTTATTTGGGCGCTGCTGATTTCTTGGGCTGTTGACC 720
 DB 277489 GTGGTAATTCAGGTCAGGCTTATTTGGGCGCTGCTGATTTCTTGGGCTGTTGACC 277548
 QY 721 ATCCGTAATTCCTTCTGGAAGAGCTGCTAATGACATGATTCCTGCTATTCCTG 780
 DB 277549 ATCCGTAATTCCTTCTGGAAGAGCTGCTAATGACATGATTCCTGCTATTCCTG 277608
 QY 781 TTGTTGTCAGATGTCAGTCAATTCCTTTCGCTCCGCGGCTTCCTTTCATTC 840
 DB 277609 TTGTTGTCAGATGTCAGTCAATTCCTTTCGCTCCGCGGCTTCCTTTCATTC 277668
 QY 841 TTAAGCCCTTACGTGATCAATTTGTGGCAAAATGGCGATGTGGATGCGAGAGAA 900
 DB 277669 TTAAGCCCTTACGTGATCAATTTGTGGCAAAATGGCGATGTGGATGCGAGAGAA 277728
 QY 901 TCCCTGCAATTTTGAACCGCATCAACGTTTCGAGCAATGCGATTCGAAAGCTTAA 960
 DB 277729 TCCCTGCAATTTTGAACCGCATCAACGTTTCGAGCAATGCGATTCGAAAGCTTAA 277788
 QY 961 CCTTTGCTCATTCGCTGCTGGGATTAAGTGTGGCGGATTCCTTTCGCGCTGACGCA 1020

DB 277789 CCTTTGCTCATTCGCTGCTGGATTCAGTGTGGCGCATTCCTTTCGCGCTGACGCA 277848
 QY 1021 CCTTTGTAAGCGTTGGCGCTGGACCATTCGTTTTCGTAATCTTTTC 1076
 DB 277849 CCTTTGTAAGCGTTGGCGCTGGACCATTCGTTTTCGTAATCTTTTC 277904

RESULT 4
 AAH21110
 ID AAH21110 standard; DNA; 753 BP.
 XX
 XX AAH21110;
 AC
 XX
 XX 05-SEP-2001 (first entry)
 DT
 XX
 XX C. glutamicum bnf DNA.
 DE
 XX
 XX L-amino acid production; bnf, bnfE; branched-chain amino acid;
 KM coryneform bacterium; leucine; isoleucine; valine; medicine;
 KM animal nutrition; ds.
 XX
 OS Corynebacterium glutamicum.
 XX
 XX EPI096010-A1.
 FN
 XX
 PD 02-MAY-2001.
 XX
 PF 11-OCT-2000; 2000EP-00122057.
 XX
 PR 27-OCT-1999; 99DE-01051708.
 XX
 PA (DEGS) DEGUSSA AG.
 PA (XERJ) FORSCHUNGSZENTRUM JUELICH GMBH.
 XX
 PI Kennerknecht N, Eggeling L, Sahn H, Pfefferle W;
 DR WPI: 2001-381595/42.
 DR P-PSDB; AAB86247.
 XX
 PT New export genes from coryneform bacteria, useful for increasing
 PT fermentative production of branched-chain amino acids.
 XX
 PS Claim 5; Page 14-15; 23pp; German.

XX This invention describes a novel isolated polynucleotide (I) containing
 CC at least one sequence that (i) is 70% identical with a sequence that
 CC encodes at least one of 251 amino acid (aa) (3) or 108 aa (5)
 CC polypeptides; (ii) encodes a polypeptide at least 70% identical with (3)
 CC or (5); (iii) is the complement of (i) or (ii), or (iv) contains at least
 CC 15 consecutive bases from (i)-(iii). The invention also describes (a)
 CC protein derived from sequences of 753 bp (2) or 324 bp (4); (b)
 CC coryneform microorganisms, especially Corynebacterium, transformed with
 CC one or more (I), where these are replicative DNA; (c) production of
 CC branched-chain L-aa by fermentation of coryneform bacteria in which the
 CC bnf and/or bnf genes (or equivalent sequences) are amplified,
 CC especially overexpressed; and (d) method for isolating the bnf and/or
 CC bnf genes. (I) is used for transformation of coryneform bacteria being
 CC used for fermentative production of branched-chain amino acids,
 CC specifically leucine, isoleucine and valine, which are useful in medicine
 CC and animal nutrition. (I) can also be used as source of primers and
 CC probes for isolation of related sequences. Transformation with (i)
 CC increases yield of branched-chain amino acids. This sequence encodes the
 CC Corynebacterium glutamicum ATCC 14752 bnf protein described in the
 CC method of the invention
 CC
 XX
 SQ Sequence 753 BP; 140 A; 211 C; 194 G; 208 T; 0 U; 0 Other;
 QY Query Match 70.0%; Score 753; DB 4; Length 753;
 Best Local Similarity 100.0%; Pred. No. 8; 1e-227;
 Matches 753; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTGCAAAAACGCAAGATTCATTCAGCCTGAGAGGTGTCGATCCAGAGCCGCTG 60

CC at least one sequence that (i) is 70% identical with a sequence that
CC encodes at least one of 251 amino acid (aa) (3) or 108 aa (5)
CC polypeptides; (ii) encodes a polypeptide at least 70% identical with (3)
CC or (5); (iii) is the complement of (i) or (ii), or (iv) contains at least
CC 15 consecutive bases from (i) - (iii). The invention also describes (a)
CC protein derived from sequences of 753 bp (2) or 324 bp (4); (b)
CC coryneform microorganisms, especially Corynebacterium, transformed with
CC one or more (i), where these are replicative DNA; (c) production of
CC branched-chain L-aa by fermentation of coryneform bacteria in which the
CC brnB and/or brnF genes (or equivalent sequences) are amplified,
CC especially overexpressed; and (d) method for isolating the brnB and/or
CC brnF genes. (1) is used for transformation of coryneform bacteria being
CC used for fermentative production of branched-chain amino acids,
CC specifically leucine, isoleucine and valine, which are useful in medicine
CC and animal nutrition. (1) can also be used as source of primers and
CC probes for isolation of related sequences. Transformation with (1)
CC increases yield of branched-chain amino acids. This sequence encodes the
CC Corynebacterium glutamicum ATCC 14752 brnB protein described in the
CC method of the invention

Sequence 324 BP; 60 A; 79 C; 80 G; 105 T; 0 U; 0 Other;

Query Match 30.1%; Score 324; DB 4; Length 324;
Best Local Similarity 100.0%; Pred. No. 1.5e-91;
Matches 324; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 753 ATGACACGATTTCTCCGTATTCCTGTTGTCGAGATGTCAGTACTT 812
DB 1 ATGACACGATTTCTCCGTATTCCTGTTGTCGAGATGTCAGTACTT 60
QY 813 GCGCTCCGGCGGTCCTGTTTAACTCTTAAGCCCTTACGTAATCAATTG 872
DB 61 GCGCTCCGGCGGTCCTGTTTAACTCTTAAGCCCTTACGTAATCAATTG 120
QY 873 AAAATGCGATGTGATGTCGACAGAAATCCTTGCCATTGACGCAATCAAGTT 932
DB 121 AAAATGCGATGTGATGTCGACAGAAATCCTTGCCATTGACGCAATCAAGTT 180
QY 933 AGCAATGCGATGTGATGTCGACAGAAATCCTTGCCATTGACGCAATCAAGTT 992
DB 181 AGCAATGCGATGTGATGTCGACAGAAATCCTTGCCATTGACGCAATCAAGTT 240
QY 993 GTGGCGCATCTTCTGGCGGTGACGCACTTGTGAGCGTTGGCGTGGACCAATCGTT 1052
DB 241 GTGGCGCATCTTCTGGCGGTGACGCACTTGTGAGCGTTGGCGTGGACCAATCGTT 300
QY 1053 TTGTTGACGTGTGATCTTTTC 1076
DB 301 TTGTTGACGTGTGATCTTTTC 324

RESULT 8

AAH65255
ID AAH65255 standard; DNA; 324 BP.

XX AAH65255;

DT 26-SEP-2001 (first entry)

XX C glutamicum coding sequence fragment SEQ ID NO: 290.

XX Corynebacterium; amino acid synthesis; vitamin; saccharide;

KM organic acid synthesis; ds.

XX Corynebacterium glutamicum.

OS EP108790-A2.

PN 20-JUN-2001.

PD 18-DEC-2000; 2000EP-00127688.

PF 16-DEC-1999; 99JP-00377484.

PR 07-APR-2000; 2000JP-00159162.
PR 03-AUG-2000; 2000JP-00280988.
XX (KYOW) KYOWA HAKKO KOGYO KK.
XX Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;
PI Tateishi N, Senoh A, Ikeda M, Ozaki A;
XX WPI: 2001-376931/40.
DR P-PSDB; AAG90036.
XX Novel polynucleotides derived from Corynebacterium, for identifying
PT mutation point of a gene, measuring expression of a gene, analyzing
PR expression profile or pattern of a gene and identifying homologous gene.
XX Claim 8; SEQ ID NO 290; 246pp + Sequence Listing; English.

CC The present invention provides a number of nucleotide and protein
CC sequences from the Corynebacterium glutamicum. These
CC are useful for identifying the mutation point of a gene derived from a
CC mutant of corynebacterium, measuring expression amount and analyzing
CC the expression profile or expression pattern of a gene derived from
CC Corynebacterium, and identifying a homologue of a gene derived from
CC Corynebacterium. Corynebacterium are useful for producing amino
CC acids, nucleic acids, vitamins, saccharides and organic acids,
CC particularly L-lysine. The present sequence is a nucleic acid described
CC in the exemplification of the invention. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from the European Patent Office

Sequence 324 BP; 60 A; 79 C; 80 G; 105 T; 0 U; 0 Other;

Query Match 30.1%; Score 324; DB 5; Length 324;
Best Local Similarity 100.0%; Pred. No. 1.5e-91;
Matches 324; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 753 ATGACACGATTTCTCCGTATTCCTGTTGTCGAGATGTCAGTACTT 812
DB 1 ATGACACGATTTCTCCGTATTCCTGTTGTCGAGATGTCAGTACTT 60
QY 813 GCGCTCCGGCGGTCCTGTTTAACTCTTAAGCCCTTACGTAATCAATTG 872
DB 61 GCGCTCCGGCGGTCCTGTTTAACTCTTAAGCCCTTACGTAATCAATTG 120
QY 873 AAAATGCGATGTGATGTCGACAGAAATCCTTGCCATTGACGCAATCAAGTT 932
DB 121 AAAATGCGATGTGATGTCGACAGAAATCCTTGCCATTGACGCAATCAAGTT 180
QY 181 AGCAATGCGATGTGATGTCGACAGAAATCCTTGCCATTGACGCAATCAAGTT 240
QY 241 GTGGCGCATCTTCTGGCGGTGACGCACTTGTGAGCGTTGGCGTGGACCAATCGTT 300
QY 301 TTGTTGACGTGTGATCTTTTC 324

RESULT 9

ACA01970
ID ACA01970 standard; DNA; 324 BP.

XX ACA01970;

DT 04-JUN-2003 (first entry)

XX C. glutamicum derived ORF SEQ ID 1961.

XX Corynebacterium; nucleic acid array; fermentation; culture; ds.

OS Corynebacterium glutamicum.
 XX DE10128510-A1.
 XX
 XX 19-DEC-2002.
 XX
 XX 13-JUN-2001; 2001DE-01028510.
 XX
 XX 13-JUN-2001; 2001DE-01028510.
 XX
 XX (DEGS) DEGUSSA AG.
 XX
 XX Farwick M, Moeckel B, Pfefferle W, Bathe B, Huthmacher K;
 XX WPI; 2003-279970/28.
 XX
 XX New nucleic acid array useful for monitoring mRNA expression of
 PT Corynebacterium glutamicum during fermentation, comprising nucleic acid
 PT from Corynebacterium glutamicum.
 XX
 XX Claim 1; Page 647; 709pp; German.
 XX
 XX This invention describes a novel nucleic acid array involving
 CC Corynebacterium glutamicum polynucleotides. The arrays are used to
 CC analyze C. glutamicum, particularly for monitoring a fermentation process
 CC to determine expression levels of C. glutamicum cellular mRNA. Such
 CC monitoring particularly differentiates between expression levels of
 CC different strains of C. glutamicum and allows the adjustment of different
 CC culture and fermentation conditions. ACA00010-ACA02188 represent C.
 CC glutamicum derived polynucleotides described in the disclosure of the
 CC invention
 XX
 XX Sequence 324 BP; 60 A; 79 C; 80 G; 105 T; 0 U; 0 Other;
 SQ
 Query Match 30.1%; Score 324; DB 7; Length 324;
 Best Local Similarity 100.0%; Pred. No. 1.5e-91;
 Matches 324; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 753 ATGACACTGATTTCTCTGATATTCCTGTTGTGACAGATGAGCAGTCACTTT 812
 DB 1 ATGACACTGATTTCTCTGATATTCCTGTTGTGACAGATGAGCAGTCACTTT 60
 QY 813 GCGCTCCGGCGGCTTCGCTTAACTTAAAGCCCTACGTAACAATTGTGCGC 872
 DB 61 GCGCTCCGGCGGCTTCGCTTAACTTAAAGCCCTACGTAACAATTGTGCGC 120
 QY 873 AAAATGCGATGTGATGCGACAGAGATCTTGCATTTTGAACCGCATCAAGTTTGC 932
 DB 121 AAAATGCGATGTGATGCGACAGAGATCTTGCATTTTGAACCGCATCAAGTTTGC 180
 QY 933 AGCAATGCGATGTGATGCGACAGATCTTGCATTTTGAACCGCATCAAGTTTGC 992
 DB 181 AGCAATGCGATGTGATGCGACAGATCTTGCATTTTGAACCGCATCAAGTTTGC 240
 QY 993 GTGGCGATTTCTTGGCGGTCGACGACCTTTGTAGCTTGGCGCTGACACATGTT 1052
 DB 241 GTGGCGATTTCTTGGCGGTCGACGACCTTTGTAGCTTGGCGCTGACACATGTT 300
 QY 1053 TTTGTGGACTGTGAATCTTTTC 1076
 DB 301 TTTGTGGACTGTGAATCTTTTC 324

RESULT 10

AAAF61693/c
 ID AAAF61693 standard; DNA; 397 BP.

XX
 AC AAAF61693;
 XX
 DT 12-JUL-2001 (first entry)
 XX
 DE C. glutamicum lrp encoding DNA fragment SEQ ID 7.
 XX

KW Fermentation; L-lysine acid production; lrp gene; lysine; isoleucine;
 KW medicine; animal feed supplement; ds.
 XX
 XX Corynebacterium glutamicum.
 XX
 XX EPI090993-A1.
 XX
 XX 11-APR-2001.
 XX
 XX 29-SEP-2000; 2000EP-00121159.
 XX
 XX 05-OCT-1999; 99DE-01047792.
 XX
 XX (DEGS) DEGUSSA-HUELS AG.
 XX
 XX Moeckel B, Pfefferle W, Puehler A, Kalinowski J, Bathe B;
 XX WPI; 2001-292927/31.
 XX
 XX New lrp gene from corynebacterium bacteria, used to prepare transformants with
 PT increased synthesis of amino acids, particularly lysine and isoleucine.
 PT
 XX
 XX Example 3; Page 16; 22pp; German.
 XX
 XX This invention describes a novel isolated nucleic acid (I) from
 CC corynebacterium bacteria used for the fermentative production of selected L-
 CC amino acids, by fermenting the amino acid-producing corynebacterium in which
 CC at least the lrp gene has been weakened or amplified, then isolating
 CC amino acids that have accumulated in the medium or cells. (I) is used to
 CC transform corynebacterium for production of L-amino acids, specifically
 CC lysine and isoleucine, which are used in medicine and particularly as
 CC animal feed supplement. It may also be used as probes and primers for
 CC isolating related sequences. Regulating expression of (I) improves
 CC production of amino acids, especially of L-lysine. This sequence encodes
 CC a fragment of the Corynebacterium glutamicum lrp protein which is used in
 CC the method described in the invention
 XX
 XX Sequence 397 BP; 96 A; 100 C; 104 G; 97 T; 0 U; 0 Other;
 SQ
 Query Match 25.4%; Score 273.6; DB 4; Length 397;
 Best Local Similarity 98.6%; Pred. No. 1.4e-75;
 Matches 276; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 QY 1 GTGAAAAACGCAAGATTCATTCAGGCTGAGGTGCGCATCCAGGAGCCCTG 60
 DB 284 GTGAAAAACGCAAGATTCATTCAGGCTGAGGTGCGCATCCAGGAGCCCTG 225
 QY 61 GACCAAGTGAATAAGTTATCGGCGCTACGAAATCGCGCAAGTCTAAAAACCTCCCTT 120
 DB 224 GACCAAGTGAATAAGTTATCGGCGCTACGAAATCGCGCAAGTCTAAAAACCTCCCTT 165
 QY 121 GCTGCAAGTTGGGCAATGACCGGATGTGATGCGTTGCTCTGTTATCAATAC 180
 DB 164 GCTGCAAGTTGGGCAATGACCGGATGTGATGCGTTGCTCTGTTATCAATAC 105
 QY 181 GGTACGATGATGAGGAGGAGCCCACTGTTTCCGCGCTGATTTGCGGAGCTCCACCGAA 240
 DB 104 GGTACGATGATGAGGAGGAGCCCACTGTTTCCGCGCTGATTTGCGGAGCTCCACCGAA 45
 QY 241 ATGCTGATCATGCGCCCTGCTGTTGTTGGGCGCAGCGCCCTG 280
 DB 44 ATGCTGATCATGCGCCCTGCTGTTGTTGGGCGCAGCGCCCTG 5

RESULT 11

AAAF61695/c
 ID AAAF61695 standard; DNA; 778 BP.

XX
 AC AAAF61695;
 XX
 DT 12-JUL-2001 (first entry)
 XX
 DE C. glutamicum lrp encoding DNA fragment SEQ ID 9.
 XX

XX Fermentation; L-amino acid production; lrp gene; lysine; isoleucine;
 KW medicine; animal feed supplement; ds.
 XX
 OS Corynebacterium glutamicum.
 XX
 PN EPI090993-A1.
 XX
 PD 11-APR-2001.
 XX
 PF 29-SEP-2000; 2000EP-00121159.
 XX
 PR 05-OCT-1999; 99DE-01047792.
 XX
 PA (DEGS) DEGUSA-HUELS AG.
 XX
 PI Moeckel B, Pfeifferle W, Fuehler A, Kalinowski J, Bathe B;
 XX
 DR WPI; 2001-292927/31.
 XX
 PT New lrp gene from coryneform bacteria, used to prepare transformants with
 PT increased synthesis of amino acids, particularly lysine and isoleucine.
 XX
 PS Example 3; Page 17; 22pp; German.
 XX
 CC This invention describes a novel isolated nucleic acid (I) from
 CC coryneform bacteria used for the fermentative production of selected L-
 CC amino acids, by fermenting the amino acid-producing coryneform in which
 CC at least the lrp gene has been weakened or amplified, then isolating
 CC amino acids that have accumulated in the medium or cells. (I) is used to
 CC transform coryneforms for production of L-amino acids, specifically
 CC lysine and isoleucine, which are used in medicine and particularly as
 CC animal feed supplement. It may also be used as probes and primers for
 CC isolating related sequences. Regulating expression of (I) improves
 CC production of amino acids, especially of L-lysine. This sequence encodes
 CC a fragment of the Corynebacterium glutamicum lrp protein which is used in
 CC the method described in the invention
 XX
 SQ Sequence 778 BP; 175 A; 208 C; 207 G; 188 T; 0 U; 0 Other;
 Query Match 25.4%; Score 273.6; DB 4; Length 778;
 Best Local Similarity 98.6%; Pred. No. 1.9e-75;
 Matches 276; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 QY 1 GTGCAAAAACGCAAGATTCATTCAAGCCTGGAGGTGTGCGCATCCAGAGCCCTG 60
 DB 284 GTGCAAAAACGCAAGATTCATTCAAGCCTGGAGGTGTGCGCATCCAGAGCCCTG 225
 QY 61 GAACCAATGATTAAGTTATCGCGCTAGAGAAATCCGCAAGGTCTAAAACTCCCTT 120
 DB 224 GAACCAATGATTAAGTTATCGCGCTAGAGAAATCCGCAAGGTCTAAAACTCCCTT 165
 QY 121 GCTGCAAGTTGGGCAATGACCCGATTGATTCGTTGGTCTCTTGTATTCAATAC 180
 DB 164 GCTGCAAGTTGGGCAATGACCCGATTGATTCGTTGGTCTCTTGTATTCAATAC 105
 QY 181 GGTACGAATGTGTGGGCAAGCCCACTGTTTTCGCGCTGATTTTCGGGGCTCAACGAA 240
 DB 104 GGTACGAATGTGTGGGCAAGCCCACTGTTTTCGCGCTGATTTTCGGGGCTCAACGAA 45
 QY 241 ATGCTGATCATCGCCCTGTTGTGGGCGACGCGCCCTGG 280
 DB 44 ATGCTGATCATCGCCCTGTTGTGGGCGACGCGCCCTGG 5
 RESULT 12
 AAH68373
 ID AAH68373 standard; DNA; 711 BP.
 XX
 AC AAH68373;
 XX
 DT 26-SEP-2001 (first entry)
 XX

DE C glutamicum coding sequence fragment SEQ ID NO: 3408.
 XX
 KW Coryneform bacterium; amino acid synthesis; vitamin; saccharide;
 XX organic acid synthesis; ds.
 XX
 OS Corynebacterium glutamicum.
 XX
 PN EPI108790-A2.
 XX
 PD 20-JUN-2001.
 XX
 PF 18-DEC-2000; 2000EP-00127688.
 XX
 PR 16-DEC-1999; 99JP-00377484.
 PR 07-APR-2000; 2000JP-00159162.
 PR 03-AUG-2000; 2000JP-00280988.
 XX
 PA (KYOW) KYOWA HAKKO KOGYO KK.
 XX
 PI Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;
 PI Tateishi N, Senoh A, Ikeda M, Ozaki A.
 XX
 DR WPI; 2001-376931/40.
 DR P-PSDB; AAG93154.
 XX
 PT Novel polynucleotides derived from Coryneform bacteria, for identifying
 PT mutation point of a gene, measuring expression of a gene, analyzing
 PT expression profile or pattern of a gene and identifying homologous gene.
 XX
 PS Claim 8; SEQ ID NO 3408; 246bp + Sequence Listing; English.
 XX
 CC The present invention provides a number of nucleotide and protein
 CC sequences from the Coryneform bacterium Corynebacterium glutamicum. These
 CC are useful for identifying the mutation point of a gene derived from a
 CC mutant of coryneform bacterium, measuring expression amount and analyzing
 CC the expression profile or expression pattern of a gene derived from
 CC Coryneform bacterium, and identifying a homologue of a gene derived from
 CC coryneform bacterium. Coryneform bacteria are useful for producing amino
 CC acids, nucleic acids, vitamins, saccharides and organic acids,
 CC particularly L-lysine. The present sequence is a nucleic acid described
 CC in the exemplification of the invention. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from the European Patent Office
 XX
 SQ Sequence 711 BP; 123 A; 190 C; 198 G; 200 T; 0 U; 0 Other;
 Query Match 11.8%; Score 127.4; DB 5; Length 711;
 Best Local Similarity 51.5%; Pred. No. 2.7e-29;
 Matches 318; Conservative 0; Mismatches 296; Indels 3; Gaps 1;
 QY 122 GTGCAAGTTGGGCAATGACCCGATTGATTCGTTGGTCTCTTGTATTCAATAC 181
 DB 53 CGATGCGTTGGGCGGTGATTCGCTGGGTTGGCCTTTGGGCTGTGTGATGTCAGACAG 112
 QY 182 GGTACGAATGTGTGGGCAAGCCCACTGTTTTCGCGCTGATTTTCGGGGCTCAACGAA 241
 DB 113 GTTTCGCTGTGTGTGGGCAAGCCCACTGTTTTCGCTGTGTGTGTGTGTGTGTGTGT 172
 QY 242 TGT 301
 DB 173 TTTGT 232
 QY 302 TGT 361
 DB 233 TGT 292
 QY 362 ACCCAATGCGCGCTTCTATTGATTTTCGCGCTTATCGAGCAAGCTTACGATCACTG 421
 DB 293 CCGGCGCGCGCGCGCTTATCGAGCTTACGAGCAAGCTTACGATCACTGATGAT 352
 QY 422 CCGGCGCGCGCGCGCTTATCGAGCTTACGAGCTTACGATCACTGATGATGATGATG 478
 DB 353 CAGCGCGCGCGCGCTTATCGAGCTTACGAGCTTACGATCACTGATGATGATGATGATG 412

QY 479 ACTCTACTGGGATTTGCGCGCTCTCACCGAGTGGCGANTCGAGATTGATTCCTTTTG 538
 Db 413 AAGCTCTGTGGGATATCCAGGAATTAATTGGCGCTTGTTGTGCAAGTGTGCGCATG 472
 QY 539 AATTAAAGGCGCTGAGTTGCGCCCTTGTCTCTCTTTGTGACGCTGATTTGATTCCT 598
 Db 473 ATCTAAAGGATGATTTTGGCCCTGACCGCGCTGTTGTGCTGCTGGCGGAGGAT 532
 QY 599 GCCGAACGAAAAAGCAGATCCCTTCTCTGCTGCTGCGAGTTTGAGCTTCAACATTCCTC 658
 Db 533 TCAAAAATACAGAGATTAATTCGTCATTAATTCGCGATGATGAGCTCTGGTTCCG 592
 QY 659 TTGTGTAATTCGAGGAGGCGCCATTTGCGGCGCTGCTGATCTTCTGGGCTGTGA 718
 Db 593 GTTTGTGGCGCCGACGAGATGCTGTATGCTTTGACACGACTTTTATCTTC 652
 QY 719 CCATCCGCTACTTCTTC 735
 Db 653 TTCTCCGCTCCGCTTC 669

RESULT 13

ACA01294
 ID ACA01294 standard; DNA; 711 BP.

ACA01294;

DT 03-JUN-2003 (first entry)

DE C. glutamicum derived ORF SEQ ID 1285.

XX Coryneform; nucleic acid array; fermentation; culture; ds.

XX Corynebacterium glutamicum.

PN DE10128510-A1.

PD 19-DEC-2002.

PF 13-JUN-2001; 2001DE-01028510.

PR 13-JUN-2001; 2001DE-01028510.

PA (DEGUSA AG.

PI Farwick M, Moeckel B, Pfeifferle W, Bathe B, Huchmacher K;

DR WPI; 2003-279970/28.

PT New nucleic acid array useful for monitoring mRNA expression of

PT Corynebacterium glutamicum during fermentation, comprising nucleic acid

PT from Corynebacterium glutamicum.

CC Claim 1; Page 456; 709pp; German.

CC This invention describes a novel nucleic acid array involving
 CC Corynebacterium glutamicum polynucleotides. The arrays are used to
 CC analyse C. glutamicum, particularly for monitoring a fermentation process
 CC to determine expression levels of C. glutamicum cellular mRNA. Such
 CC monitoring particularly differentiates between expression levels of
 CC different strains of C. glutamicum and allows the adjustment of different
 CC culture and fermentation conditions. ACA00010-ACA02188 represent C.
 CC glutamicum derived polynucleotides described in the disclosure of the
 CC invention

XX Sequence 711 BP; 123 A; 190 C; 198 G; 200 T; 0 U; 0 Other;

XX Query Match 11.8%; Score 127.4; DB 7; Length 711;

XX Best Local Similarity 51.5%; Pred. No. 2.7e-29;

XX Matches 318; Conservative 0; Mismatches 296; Indels 3; Gaps 1;

QY 122 CTGCAAGTGTGGGCAATGACCGGATTTGGTATTCGTTGCTCTTGGTATTCATACG 181

Db 53 CGGTGCGTTTGGAGGCTGATTCGCTGGGTTTGCCCTTTGGGCTGTGATGATCCAGACG 112
 QY 182 GCTACGATGATGAGGACAGCCCACTGTTTTCGGGCTGATTTTCGGGGCTCCACCGAA 241
 Db 113 GTTTCGCTGTGTGAGAGAGCCGATTTTCTCTTCGTGATCTATGCCGTTGATGAAAT 172
 QY 242 TGTGTGTAATGCGCCCTGCTGTGTGAGGCGAGGCGCCCTGGGGCCATTCGGGCTACCAAT 301
 Db 173 TTCTGCAATGCGCATGATGACCGAGATGAGCCCTGTTTTCGGCGGCTGCTGTGTT 232
 QY 302 TGTGTGTAATTCGCGCAGATTTCTATGCGTTTCAATTCGCGCTGATGATGATGATGAT 361
 Db 233 TCAATGTAATTTCCGCCATTTTCTAGGCTCACTCCACGCGCAACGCAATCAAGT 292
 QY 362 ACCCATTCGCGCTTTCTATGCTTTTCGGCTTTCGCTTATGAGAGAGCTTACGCACTG 421
 Db 293 CGGCGCGCGCGCGCTTATTCACCTTACGCGCTTACGAGATCTTACGCGCATGCTGT 352
 QY 422 CGGCGAGCGCGCGAGGCTGTGCGGTG--GCGATTATCTCAATGCAATGAGCTTTC 478
 Db 353 CAGCCGCCACCTGGCGATATCAGGAGAGCGGGGTGCTTACGTTCAAAATTTGTGCG 412
 QY 479 ACTCTACTGGGATTTGCGCGCTTCAACGAGTGGCCATGCGAAGTTGATTCCTTTTG 538
 Db 413 AAGCTCTGTGGGATTAATCCAGGAATTAATTGGCGCTTGTTGTGCAAGTGTGCGCATG 472
 QY 539 AATTAAAGGCGCTGAGTTGCGCCCTTGTCTCTCTTTGTGACGCTGATTTGATTCCT 598
 Db 473 ATCTAAAGGATGATTTTGGCCCTGACCGCGCTGTTGTGCTGCTGGCGGAGGAT 532
 QY 599 GCCGAACGAAAAAGCAGATCCCTTCTCTGCTGCTGCGAGTTTGAGCTTCAACATTCCTC 658
 Db 533 TCAAAAATACAGAGATTAATTCGTCATTAATTCGCGATGATGAGCTCTGGTTTCG 592
 QY 659 TTGTGTAATTCGAGGAGGCGCCATTTGCGGCGCTGCTGATCTTCTGGTGTGTTGA 718
 Db 593 GTTTGTGGCGCCGACGAGATGCTGTATGCTTTGACCAAGTATTTTGAATCTTC 652
 QY 719 CCATCCGCTACTTCTTC 735
 Db 653 TTCTCCGCTCCGCTTC 669

RESULT 14

AAH68534/c
 ID AAH68534 standard; DNA; 309400 BP.

XX AAH68534;

DT 26-SEP-2001 (first entry)

DE C glutamicum coding sequence fragment SEQ ID NO: 7069.

XX Coryneform bacterium; amino acid synthesis; vitamin; saccharide;

XX Corynebacterium glutamicum.

PN EP1108790-A2.

PD 20-JUN-2001.

PF 18-DEC-2000; 2000EP-00127688.

PR 16-DEC-1999; 99JP-00377484.

PR 07-APR-2000; 2000JP-00159162.

PR 03-AUG-2000; 2000JP-00280988.

PA (KYOWA) KYOWA HAKKO KOGYO KK.

PI Nakagawa S, Mizoguchi H, Ando S, Hayaishi M, Ochiai K, Yokoi H;

PI Tateishi N, Senoh A, Ikeda M, Ozaki A;

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 19, 2004, 10:57:33 ; Search time 90.5968 seconds

(without alignments)
6590.898 Million cell updates/sec

Title: US-10-608-504-1_COPY_101_1176
Perfect score: 1076
Sequence: 1 gtgcataaagcgaagat.....ctgactgtgtaatttc 1076

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 662709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA.*
1: /cgn2_6/ptodata/2/ina/5A.COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B.COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A.COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B.COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PTCUS.COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfile1.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1076	100.0	1271	4	US-09-471-803A-1
2	1076	100.0	1271	4	US-09-471-803A-6
3	753	70.0	753	4	US-09-471-803A-2
4	324	30.1	324	4	US-09-471-803A-4
5	58.6	5.4	756	4	US-09-469-039A-3016
6	52.4	4.9	7218	1	US-08-232-443-14
7	39.8	3.7	768	4	US-09-107-532A-1266
8	39.2	3.6	1086	4	US-09-252-991A-11376
9	39.2	3.6	1086	4	US-09-252-991A-11418
10	39	3.6	399	4	US-09-621-976-8976
11	37	3.4	474	4	US-09-621-976-18033
12	36.8	3.4	1140	4	US-09-252-991A-3205
13	36.8	3.4	2061	4	US-09-252-991A-2761
14	36.4	3.4	751	3	US-09-020-956-12
15	36.4	3.4	751	3	US-09-030-607-12
16	36.4	3.4	751	4	US-09-439-313-12
17	36.4	3.4	751	4	US-09-352-616A-12
18	36.4	3.4	751	4	US-09-233-149A-12
19	36.4	3.4	751	4	US-09-159-812-12
20	36.4	3.4	751	4	US-09-636-215-12
21	36.4	3.4	751	4	US-09-685-166A-12
22	36.4	3.4	751	4	US-09-115-453-12
23	36.4	3.4	751	4	US-09-688-489-12
24	35.4	3.3	430	4	US-09-621-976-16656
25	35.2	3.3	801	4	US-09-328-352-2399
26	35.2	3.3	4403765	3	US-09-103-840A-2
27	35.2	3.3	4411529	3	US-09-103-840A-1

28	34.8	3.2	867	4	US-09-489-039A-5047	Sequence 5047, Ap
29	34.6	3.2	666	2	US-08-875-034A-1	Sequence 1, Appl
30	34.4	3.2	744	4	US-09-489-039A-1307	Sequence 1307, Ap
31	34.4	3.2	1224	4	US-09-266-965-22	Sequence 22, Appl
32	34.4	3.2	1224	4	US-09-266-965-74	Sequence 74, Appl
33	34.4	3.2	18331	4	US-09-266-965-96	Sequence 96, Appl
34	34.2	3.2	594	4	US-09-252-991A-13729	Sequence 13729, A
35	34.2	3.2	678	4	US-09-252-991A-13832	Sequence 13832, A
36	33.4	3.1	1875	3	US-08-433-522A-1	Sequence 21, Appl
37	33.4	3.1	2949	3	US-08-433-522A-1	Sequence 1, Appl
38	33.4	3.1	2849	3	US-09-135-166-1	Sequence 1, Appl
39	33.4	3.1	2849	3	US-08-942-046-1	Sequence 1, Appl
40	33.4	3.1	2950	3	US-08-942-046-5	Sequence 5, Appl
41	33.4	3.1	2950	3	US-09-135-166-5	Sequence 5, Appl
42	33.4	3.1	2950	3	US-08-942-046-5	Sequence 5, Appl
43	33.4	3.1	2984	3	US-08-433-522A-3	Sequence 3, Appl
44	33.4	3.1	2984	3	US-09-135-166-3	Sequence 3, Appl
45	33.4	3.1	2984	3	US-08-942-046-3	Sequence 3, Appl

ALIGNMENTS

RESULT 1
US-09-471-803A-1
Sequence 1, Application US/09471803A
Patent No. 6613545
GENERAL INFORMATION:
APPLICANT: KENNERKNECHT, NICOLE
APPLICANT: SAHM, HERMANN
APPLICANT: EGGERLING, LOTAR
APPLICANT: PFEFFERLE, WALTER
TITLE OF INVENTION: NUCLEOTIDE SEQUENCE CODING FOR THE EXPORT OF
TITLE OF INVENTION: BRANCHED CHAIN AMINO ACIDS, PROCESS FOR THE
FILE REFERENCE: 21123/265496/MAS
CURRENT APPLICATION NUMBER: US/09/471, 803A
CURRENT FILING DATE: 1999-12-23
PRIOR APPLICATION NUMBER: DE 199 51 708.8
PRIOR FILING DATE: 1999-10-27
NUMBER OF SEQ ID NOS: 12
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 1
LENGTH: 1271
TYPE: DNA
ORGANISM: Corynebacterium glutamicum
FEATURE:
NAME/KEY: gene
LOCATION: (101)..(853)
OTHER INFORMATION: brnF
NAME/KEY: gene
LOCATION: (853)..(1176)
OTHER INFORMATION: brnE
OTHER INFORMATION: ATCC14752
US-09-471-803A-1
Query Match 100.0%; Score 1076; DB 4; Length 1271;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1076; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GTGCAAAAACGCAAGATTCATTCAGCTCGAGGTTCGCCATCCAGGACGCCCTG 60
DB 101 GTGCAAAAACGCAAGATTCATTCAGCTCGAGGTTCGCCATCCAGGACGCCCTG 160
QY 61 GAACGAGATTAAGGTATCGCGCTGCAAGATTCGCAAGTCTAAACCTCCCTT 120
DB 161 GAACGAGATTAAGGTATCGCGCTGCAAGATTCGCAAGTCTAAACCTCCCTT 220
QY 121 GCTGCAAGTTTGCGCATGATCCGATGATGATGCTTGTCTCTGTTATTCATTC 180
DB 221 GCTGCAAGTTTGCGCATGATCCGATGATGATGCTTGTCTCTGTTATTCATTC 280
QY 181 GCTGCAAGTTTGCGCATGATCCGATGATGATGCTTGTCTCTGTTATTCATTC 240

```

Db 281 GGCACAGATGATGGGAGGCCCACTGTTTTCGGGCGCTGATTTTCGGGGGCTCCACGGA 340
Qy 241 ATGCTGATCATGCGCTGTTGTGTGGGCGAGCGCCCTCGGGCGCATCGGCTCACCACA 300
Db 341 ATGCTGATCATGCGCTGTTGTGTGGGCGAGCGCCCTCGGGCGCATCGGCTCACCACA 400
Qy 301 TTGCTGATGAACTTCCGCGCATGATTTCTATGCTTTTCATTCGCGCTGATGATGATG 360
Db 401 TTGCTGATGAACTTCCGCGCATGATTTCTATGCTTTTCATTCGCGCTGATGATGATG 460
Qy 361 AACCCATTGCGCGCTTCTATTCGTTTTCGCGCTTATCGAGAGCTTACGAGTCACT 420
Db 461 AACCCATTGCGCGCTTCTATTCGTTTTCGCGCTTATCGAGAGCTTACGAGTCACT 520
Qy 421 GCGGCGAGCGCGCGAGCTGATCGCGCGAGGAGCTTATCTCAATGCAATAGAGTTTAC 480
Db 521 GCGGCGAGCGCGCGAGCTGATCGCGCGAGGAGCTTATCTCAATGCAATAGAGTTTAC 580
Qy 481 TCCTACTGAGTATTCGCGGATCTCAACGAGTGGCGATCGAGAGTTGATTCCTTTGAA 540
Db 581 TCCTACTGAGTATTCGCGGATCTCAACGAGTGGCGATCGAGAGTTGATTCCTTTGAA 640
Qy 541 ATTAAAGGCTCTGAGTTGGCTTTGCTCTCTTTTGTGAGCTGATCTTTGATTCCTG 600
Db 641 ATTAAAGGCTCTGAGTTGGCTTTGCTCTCTTTTGTGAGCTGATCTTTGATTCCTG 700
Qy 601 CGAAGGAAAGAGAGATCCCTTCTCTGCTGCTGAGGTTGAGCTTCAACATTTGCTT 660
Db 701 CGAAGGAAAGAGAGATCCCTTCTCTGCTGCTGAGGTTGAGCTTCAACATTTGCTT 760
Qy 661 GTGTAATTCAGAGTCAGGCTTATTTGGGCGCTGCTGATCTTTGAGGCTTGTAC 720
Db 761 GTGTAATTCAGAGTCAGGCTTATTTGGGCGCTGCTGATCTTTGAGGCTTGTAC 820
Qy 721 ATCCGATCTTCTTTGGGAAAGGCTCTAATATGACAGATGATTTCTCGTATTCCT 780
Db 821 ATCCGATCTTCTTTGGGAAAGGCTCTAATATGACAGATGATTTCTCGTATTCCT 880
Qy 781 TTGTTGTCAGATATGTCAGTCACTTTTGGGCTCGGGCGGTTCCGTTTAAATCC 840
Db 881 TTGTTGTCAGATATGTCAGTCACTTTTGGGCTCGGGCGGTTCCGTTTAAATCC 940
Qy 841 TTAAGCCCTCAGTGAATCACAATTTGTGGGCAAAATGGCATGTGATGTCAGACAGAA 900
Db 941 TTAAGCCCTCAGTGAATCACAATTTGTGGGCAAAATGGCATGTGATGTCAGACAGAA 1000
Qy 901 TCCTTGCAATTTTACGCGCATCAACGTTTGCAGCAATGCGATGATGTAAGACTTAA 960
Db 1001 TCCTTGCAATTTTACGCGCATCAACGTTTGCAGCAATGCGATGATGTAAGACTTAA 1060
Qy 961 CCTTGATCTATTTGCGGTTGATTAAGATGTGGGCGATCTTCTTGGGCGTGAAGCA 1020
Db 1061 CCTTGATCTATTTGCGGTTGATTAAGATGTGGGCGATCTTCTTGGGCGTGAAGCA 1120
Qy 1021 CCTTGATCTATTTGCGGCTGAGCAACATGTTTGTGGAAGTGTGATCTTTTC 1076
Db 1121 CCTTGATCTATTTGCGGCTGAGCAACATGTTTGTGGAAGTGTGATCTTTTC 1176

```

```

RESULT 2
US-09-471-803A-6
; Sequence 6, Application US/09471803A
; Patent No. 6613545
; GENERAL INFORMATION:
; APPLICANT: KENNERKNECHT, NICOLE
; APPLICANT: SAHM, HERMANN
; APPLICANT: EGSELM, LOTHAR
; APPLICANT: PFEFFERLE, WALTER
; TITLE OF INVENTION: NOCLEOTIDE SEQUENCE CODING FOR THE EXPORT OF
; TITLE OF INVENTION: BRANCHED CHAIN AMINO ACIDS, PROCESS FOR THE
; FILE REFERENCE: 21123/265496/MAS

```

```

CURRENT APPLICATION NUMBER: US/09/471, 803A
CURRENT FILING DATE: 1999-12-23
PRIOR APPLICATION NUMBER: DE 199 51 708.8
PRIOR FILING DATE: 1999-10-27
NUMBER OF SEQ ID NOS: 12
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 6
LENGTH: 1271
TYPE: DNA
ORGANISM: Corynebacterium glutamicum
FEATURE:
NAME/KEY: gene
LOCATION: (101)..(853)
OTHER INFORMATION: brnF
NAME/KEY: gene
LOCATION: (853)..(1176)
OTHER INFORMATION: brnE
OTHER INFORMATION: ATCC13032
US-09-471-803A-6

```

```

Query Match 100.0%; Score 1076; DB 4; Length 1271;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1076; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 1 GTGCAAAAAGCGAGATTCATTCAGAGCTGAGGCTGTGCGCATCCAGGAGCCCTG 60
Db 101 GTGCAAAAAGCGAGATTCATTCAGAGCTGAGGCTGTGCGCATCCAGGAGCCCTG 160
Qy 61 GAACGATGATTAAGATTCAGGCGCTACGAAATCGCGCAAGTCTAATAAATCCCTT 120
Db 161 GAACGATGATTAAGATTCAGGCGCTACGAAATCGCGCAAGTCTAATAAATCCCTT 220
Qy 121 GCTGAGGTTTGGGATGATGACCGATGATGATGATGATGATGATGATGATGATG 180
Db 221 GCTGAGGTTTGGGATGATGACCGATGATGATGATGATGATGATGATGATGATG 280
Qy 181 GGTACGAATGATGAGGAGAGCCCACTGTTTCCGAGCTGATTTGCGGCGCTCAGCGAA 240
Db 281 GGTACGAATGATGAGGAGAGCCCACTGTTTCCGAGCTGATTTGCGGCGCTCAGCGAA 340
Qy 241 ATGCTGATATGCGCTGCTGTGTGGCGAGAGCCCTGAGGCGCATGCGCTCAGCA 300
Db 341 ATGCTGATATGCGCTGCTGTGTGGCGAGAGCCCTGAGGCGCATGCGCTCAGCA 400
Qy 301 TTGCTGATGATCTCGGACGATTTATGAGTTTCAATCCGCTGATGATGATGATG 360
Db 401 TTGCTGATGATCTCGGACGATTTATGAGTTTCAATCCGCTGATGATGATGATG 460
Qy 361 AACCCATTGCGCGCTTCTATTCGTTTTCGCGCTTATCGAGAGCTTACGAGTCACT 420
Db 461 AACCCATTGCGCGCTTCTATTCGTTTTCGCGCTTATCGAGAGCTTACGAGTCACT 520
Qy 421 GCGGCGAGCGCGCGAGCTGATCGCGCGAGGAGCTTATCTCAATGCAATAGAGTTTAC 480
Db 521 GCGGCGAGCGCGCGAGCTGATCGCGCGAGGAGCTTATCTCAATGCAATAGAGTTTAC 580
Qy 481 TCCTACTGAGTATTCGCGGATCTCAACGAGTGGCGATGCGAGTGAATTCCTTTGAA 540
Db 581 TCCTACTGAGTATTCGCGGATCTCAACGAGTGGCGATGCGAGTGAATTCCTTTGAA 640
Qy 541 ATTAAAGGCTCTGAGTTGGCTTTGCTCTCTTTTGTGAGCTGATCTTTGATTCCTG 600
Db 641 ATTAAAGGCTCTGAGTTGGCTTTGCTCTCTTTTGTGAGCTGATCTTTGATTCCTG 700
Qy 601 CGAAGGAAAGAGATCCCTTCTCTGCTGCTGAGGTTGAGGTTGAGGTTGAGGTTG 660
Db 701 CGAAGGAAAGAGATCCCTTCTCTGCTGCTGAGGTTGAGGTTGAGGTTGAGGTTG 760
Qy 661 GTGTAATTCAGAGTACGCGCTTATTTGGGCGCTGATCTTTGATGATGATGATG 720
Db 761 GTGTAATTCAGAGTACGCGCTTATTTGGGCGCTGATCTTTGATGATGATGATG 820
Qy 721 ATCCGATCTTCTTTGGGAAAGGCTCTAATATGACAGATGATTTCTCGTATTCCT 780

```

```

Db      821 ATCCGTAATCTTCTTGGGAAAGGCTGCTAAATGACAATGTTCTCTGTAATCTTC 880
QY      781 TTGTTGTCAGATATGTCAGTCAATTAATTTGCGCTCCGGGCGGTTCCGTTTAAATCC 840
Db      881 TTGTTGTCAGATATGTCAGTCAATTAATTTGCGCTCCGGGCGGTTCCGTTTAAATCC 940
QY      841 TTAAAGCCCTACAGTATCAATTTGTCGCAAAATGCGATGTCGATGCGGCAAGAA 900
Db      941 TTAAAGCCCTACAGTATCAATTTGTCGCAAAATGCGATGTCGATGCGGCAAGAA 1000
QY      901 TCCATTGCAATTTGACCCGATCAACGTTTGCAGCAATGCAATGATCTGAACATCTAA 960
Db      1001 TCCATTGCAATTTGACCCGATCAACGTTTGCAGCAATGCAATGATCTGAACATCTAA 1060
QY      961 CCTTTGATCATTTGCGGCTGCGATTAAGTGTGGGCGATCTTTTGGGGTCGACGCA 1020
Db      1061 CCTTTGATCATTTGCGGCTGCGATTAAGTGTGGGCGATCTTTTGGGGTCGACGCA 1120
QY      1021 CCTTTGAGGCTTGGGCTGCGACATCGTTTGTGTAATGTAATCTTTTC 1076
Db      1121 CCTTTGAGGCTTGGGCTGCGACATCGTTTGTGTAATGTAATCTTTTC 1176

```

RESULT 3

US-09-471-803A-2

Sequence 2, Application US/09471803A

Patent No. 6613545

GENERAL INFORMATION:

APPLICANT: KENNERKNECHT, NICOLE

APPLICANT: SAHM, HERMANN

APPLICANT: EGGELING, WALTER

APPLICANT: PEEFERLE, WALTER

TITLE OF INVENTION: NUCLEOTIDE SEQUENCE CODING FOR THE EXPORT OF

TITLE OF INVENTION: BRANCHED CHAIN AMINO ACIDS, PROCESS FOR THE

TITLE OF INVENTION: ISOLATION THEREOF AND USE THEREOF

FILE REFERENCE: 21123/265496/MAS

CURRENT APPLICATION NUMBER: US/09/471,803A

CURRENT FILING DATE: 1999-12-23

PRIOR APPLICATION NUMBER: DE 199 51 708.8

PRIOR FILING DATE: 1999-10-27

NUMBER OF SEQ ID NOS: 12

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 2

LENGTH: 753

TYPE: DNA

ORGANISM: Corynebacterium glutamicum

FEATURE:

NAME/KEY: CDS

LOCATION: (1)..(753)

OTHER INFORMATION: b2mf

OTHER INFORMATION: ATCC14752

US-09-471-803A-2

Query Match 70.0%; Score 753; DB 4; Length 753;

Best Local Similarity 100.0%; Pred. No. 9.6e-236;

Matches 753; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY      1 GTGCAAAAACGCAAGATTCATTCAAGCTGAGGTGTGCGCATCCAGACGCCCTG 60
Db      1 GTGCAAAAACGCAAGATTCATTCAAGCTGAGGTGTGCGCATCCAGACGCCCTG 60
QY      61 GAACAGATGATAAGGTTATCGGCGTACAGAAATCGCGCAAGGCTTAAAAACCTCCCTT 120
Db      61 GAACAGATGATAAGGTTATCGGCGTACAGAAATCGCGCAAGGCTTAAAAACCTCCCTT 120
QY      121 GCTGAGGTTTGGCGATGACCCGATGTAATGCGTTGCTCTTGTTATTCATATAC 180
Db      121 GCTGAGGTTTGGCGATGACCCGATGTAATGCGTTGCTCTTGTTATTCATATAC 180
QY      181 GGTGACGATGTTGGGCGATGACCCGATGTAATGCGTTGCTCTTGTTATTCATATAC 240
Db      181 GGTGACGATGTTGGGCGATGACCCGATGTAATGCGTTGCTCTTGTTATTCATATAC 240

```

```

QY      241 ATGCTGTCATCGCCCTCTGTTGTGGGCGACGCCCCCTGGGCGCATGCGGCTCACCACA 300
Db      241 ATGCTGTCATCGCCCTCTGTTGTGGGCGACGCCCCCTGGGCGCATGCGGCTCACCACA 300
QY      301 TTGCTGTGAACCTTCGCGACAGATATTCATGCGTTTCAATCCGCTGCATGTGATCA 360
Db      301 TTGCTGTGAACCTTCGCGACAGATATTCATGCGTTTCAATCCGCTGCATGTGATCA 360
QY      361 AACCCTATGCCCCCTTCTATTCGCTTTTGGCGCTTATCAGCAAGCAAGCTTAGCGTCACT 420
Db      361 AACCCTATGCCCCCTTCTATTCGCTTTTGGCGCTTATCAGCAAGCAAGCTTAGCGTCACT 420
QY      421 GCGGCGACGCGCGAGGCTGTGCGGCTGCGGCACTTATCTCAATGCAATAGCGTTTAC 480
Db      421 GCGGCGACGCGCGAGGCTGTGCGGCTGCGGCACTTATCTCAATGCAATAGCGTTTAC 480
QY      481 TCCTACTGAGTATTCGCGGCTGTCAACGGAATGCGAGTGCAGATGATTCCTTTTGA 540
Db      481 TCCTACTGAGTATTCGCGGCTGTCAACGGAATGCGAGTGCAGATGATTCCTTTTGA 540
QY      541 ATTAAGGCGCTGAGTTCGCGCTTGTCTCTCTTGTCAAGCTGACTTGGATTCTGCG 600
Db      541 ATTAAGGCGCTGAGTTCGCGCTTGTCTCTCTTGTCAAGCTGACTTGGATTCTGCG 600
QY      601 CGAAGCAAAAACGATCCCTTCTGCTGCTGCTGCAAGTTTGAAGTTGACATTCCTT 660
Db      601 CGAAGCAAAAACGATCCCTTCTGCTGCTGCTGCAAGTTTGAAGTTGACATTCCTT 660
QY      661 GTGTAATTCAGGTCAGGCGCTTATTTGGGCGCTGATCTTGGTGTGTAAC 720
Db      661 GTGTAATTCAGGTCAGGCGCTTATTTGGGCGCTGATCTTGGTGTGTAAC 720
QY      721 ATCCGTAATCTTCTTGGGAAAGGCTGTAA 753
Db      721 ATCCGTAATCTTCTTGGGAAAGGCTGTAA 753

```

RESULT 4

US-09-471-803A-4

Sequence 4, Application US/09471803A

Patent No. 6613545

GENERAL INFORMATION:

APPLICANT: KENNERKNECHT, NICOLE

APPLICANT: SAHM, HERMANN

APPLICANT: EGGELING, WALTER

APPLICANT: PEEFERLE, WALTER

TITLE OF INVENTION: NUCLEOTIDE SEQUENCE CODING FOR THE EXPORT OF

TITLE OF INVENTION: BRANCHED CHAIN AMINO ACIDS, PROCESS FOR THE

TITLE OF INVENTION: ISOLATION THEREOF AND USE THEREOF

FILE REFERENCE: 21123/265496/MAS

CURRENT APPLICATION NUMBER: US/09/471,803A

CURRENT FILING DATE: 1999-12-23

PRIOR APPLICATION NUMBER: DE 199 51 708.8

PRIOR FILING DATE: 1999-10-27

NUMBER OF SEQ ID NOS: 12

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 4

LENGTH: 324

TYPE: DNA

ORGANISM: Corynebacterium glutamicum

FEATURE:

NAME/KEY: CDS

LOCATION: (1)..(324)

OTHER INFORMATION: b2mf

OTHER INFORMATION: ATCC14752

US-09-471-803A-4

Query Match 30.1%; Score 324; DB 4; Length 324;

Best Local Similarity 100.0%; Pred. No. 8.1e-96;

Matches 324; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY      753 ATGACAACTGATTTCTCCCTGTAATCTCTCTTGTGTGCGAGTATGTCAGTACTTACTTT 812

```

DB 1 ATGCAACTGATTTCTCTGATTTCTCTTTGTCGAGATGTCATTTCTTTT 60
QY 813 GCGCTCCGCGGCGTTCCGTTCTTAATCCTTAAGCCCCCTACGTGAATCAATTTGGGC 872
DB 61 GCGCTCCGCGGCGTTCCGTTCTTAATCCTTAAGCCCCCTACGTGAATCAATTTGGGC 120
QY 873 AAAATGCGATGTGATGTCGACAGAGATCCTTGGCATTTTGAACGATCAACGTTTGGC 932
DB 121 AAAATGCGATGTGATGTCGACAGAGATCCTTGGCATTTTGAACGATCAACGTTTGGC 180
QY 933 AGCATGCGATGTGATGTCGACAGATCCTTGGCATTTTGAACGATCAACGTTTGGC 992
DB 181 AGCATGCGATGTGATGTCGACAGATCCTTGGCATTTTGAACGATCAACGTTTGGC 240
QY 993 GTGGCGCATCTTTCTTGGCGGCTGACGACCTTTGTCAGCGTTGGCGCTGGACCATGTT 1052
DB 241 GTGGCGCATCTTTCTTGGCGGCTGACGACCTTTGTCAGCGTTGGCGCTGGACCATGTT 300
QY 1053 TTTGTTGACGTGATGTTCTTTTC 1076
DB 301 TTTGTTGACGTGATGTTCTTTTC 324

RESULT 5

US-09-489-039A-3016
Sequence 3016, Application US/09489039A
Patent No. 6610836
GENERAL INFORMATION:
APPLICANT: Gary Bretton et. al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
FILE REFERENCE: 2709, 2004001
CURRENT APPLICATION NUMBER: US/09/489,039A
PRIOR FILING DATE: 2000-01-27
PRIOR APPLICATION NUMBER: US 60/117,747
PRIOR FILING DATE: 1999-01-25
NUMBER OF SEQ ID NOS: 14342
SEQ ID NO 3016
LENGTH: 756
TYPE: DNA
ORGANISM: Klebsiella pneumoniae
US-09-489-039A-3016

Query Match

Best Local Similarity 52.4%; Score 58.6; DB 4; Length 756;
Matches 130; Conservative 0; Mismatches 119; Indels 0; Gaps 0;

QY 95 TCGCGGAGGTCTAAACCTCCCTGCTGACAGTTGGGATGTAACCGATGTTGTTG 154
DB 62 TCGCGGAGGTCTAAACCTCCCTGCTGACAGTTGGGATGTAACCGATGTTGTTG 121
QY 155 CGTTGGTCTCTGTTATTAATACGCTACGATGTTGGGACGCCCATGTTTCG 214
DB 122 CCTCGAGCATGATGCTACCCGCTGCTCATCACTTCCGTTGCGATTCG 181
QY 215 GCTGATTTTTCGGGCTGACGAGAAATGCTGCTACGCTGTTGGGCGGACGCG 274
DB 182 GATTAATTAACGCGCGGCGGACGATGTTGATACCCGCAATGTTGGCGCGGCAATT 241
QY 275 CCCTGGGCGCATGCGGCTACCCACATGCTGTTGTAATTCGCGCAATTAATGCGT 334
DB 242 CGCTGGGCGCATGCGGCTGCGGCTGAGGATGATGCGCATGATGCGCGCATGATGCGC 301
QY 335 TTTGATTTCC 343
DB 302 CTTCACCTGC 310

RESULT 6

US-08-232-463-14
Sequence 14, Application US/08232463
Patent No. 5670367

GENERAL INFORMATION:

APPLICANT: DORNER, F.
APPLICANT: SCHEIFLINGER, F.
APPLICANT: FALKNER, F. G.
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,463
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935,313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 IMMU
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: pTZgpt-Fls
US-08-232-463-14

Query Match

Best Local Similarity 4.9%; Score 52.4; DB 1; Length 7218;
Matches 14; Conservative 22; Mismatches 157; Indels 0; Gaps 0;

QY 460 TCAATGCAATPAGGTTTCACTCTCTAGTATTCGCGGCTCTCACCGAGTGGCGATC 519
DB 1050 TCGAGGAGGCTTGAGATTTTGGATTTTAAAGGCTGAGTTGCGCTTCTCTTTGTC 1109
QY 520 GCAGAGTTGATCCTTTTGAATTAAGGCTGAGTTGCGCTTCTCTTTGTC 579
DB 1110 TTTTGGTCTCTGTTATTAATTAAGGCTGAGTTGCGCTTCTCTTTGTC 1169
QY 580 ACGCTGACTTGAATTCGTCGACGAAAGCAATGCTCTGCTGCTGCGAGGT 639
DB 1170 TTTTGGTCTCTGTTATTAATTAAGGCTGAGTTGCGCTTCTCTTTGTC 1229
QY 640 TTAGGCTTCACTGCTCTGTTGTTATTAAGGCTGAGTTGCGCTTCTCTTTGTC 699
DB 1230 TTTTGGTCTCTGTTATTAATTAAGGCTGAGTTGCGCTTCTCTTTGTC 1289
QY 700 ATCTCTGGGCTGCTGACATCGGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 759
DB 1290 TTTTGGTCTCTGTTATTAATTAAGGCTGAGTTGCGCTTCTCTTTGTC 1349
QY 760 CTGATTTCTCTGTTATTAATTAAGGCTGAGTTGCGCTTCTCTTTGTC 819
DB 1350 TTTTGGTCTCTGTTATTAATTAAGGCTGAGTTGCGCTTCTCTTTGTC 1409

QY 820 GGCGGCTTCTTCTTAATCCTTAAGCCCTTA 851
Db 1410 YYYYYYYYYYYYYYYYYYYYGTACCA 1441

RESULT 7

US-09-107-532A-1266
Sequence 1266, Application US/09107532A
Patent No. 6583275
GENERAL INFORMATION:
APPLICANT: Lynn A Doucette-Stamm and David Bush
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
NUMBER OF SEQUENCES: 7310
CORRESPONDENCE ADDRESS:
ADDRESSEE: GENOME THERAPEUTICS CORPORATION
STREET: 100 Beaver Street
CITY: Waltham
STATE: Massachusetts
COUNTRY: USA
ZIP: 02354
COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660
COMPUTER: PC
OPERATING SYSTEM: <Unknown>
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/107,532A
FILING DATE: 30-Jun-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/085,598
FILING DATE: 14 May 1998
APPLICATION NUMBER: 60/051571
FILING DATE: July 2, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Arinlelelo, Pamela Deneke
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-012
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781)893-5007
TELEFAX: (781)893-8277
INFORMATION FOR SEQ ID NO: 1266:
SEQUENCE CHARACTERISTICS:
LENGTH: 768 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Enterococcus faecium
FEATURE:
NAME/KEY: misc feature
LOCATION: (B) LOCATION 1..768
SEQUENCE DESCRIPTION: SEQ ID NO: 1266:
US-09-107-532A-1266

Query Match 3.7%; Score 39.8; DB 4; Length 768;
Best Local Similarity 49.3%; Pred. No. 0.0097;
Matches 104; Conservative 0; Mismatches 107; Indels 0; Gaps 0;

QY 107 TAAAACTCCCTTGTGAGGTTGGGATGACCCGATGGATGGATTGGCTTCT 166
Db 92 TCAAAAGACGTTACCTACCGTTTTCGTTATATCGATTGAGCTTGCAATTGGATCG 151
QY 167 TGGTTTCAATACGCGCTACGATGATGAGGAGCCCACTGTTTCCGCGCTGATTTGCG 226
Db 152 TTGGGAAAGCTGCGGATTTCACTACGATGCTGCTGATGATGCTGCTGCTGCTATG 211
QY 227 CCGGCTCCACCGAATGCTGCTGATGCGCCCTGCTGTTGGGCGAGCGCCCTGGGCGCCA 286
Db 212 CTGGTTCTGCGCAATTATACAGTCAGATGCTGCTAGTCAACAGCCCATTTGCTTCCA 271

QY 287 TCGCGCTACCAACATTTGCTGTGAATCTCCG 317
Db 272 TCGTTTCTGACCTTTCTAGTCAATTCGCCG 302

RESULT 8

US-09-252-991A-11376
Sequence 11376, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 11376
LENGTH: 1038
TYPE: DNA
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-11376

Query Match 3.6%; Score 39.2; DB 4; Length 1038;
Best Local Similarity 55.0%; Pred. No. 0.018;
Matches 77; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

QY 209 TTTCCGCTGATTTTCGCGGCTCCACCGAATGCTGTCATCGCCCTGTTGTGCGCG 268
Db 473 TGTCCGCTATCGCTTTCGCGGCGCGCCCGCCAGTTGTGGATCGGCAATGCTCAAGGCG 532
QY 269 CAGCGCCCTTGGCGCGCCATCGCTCACCAATGCTGTGTAATCTCCGCACTATCT 328
Db 533 GTGCCGGGTTCTTTGATATCTTCAACACCTCTCTGACCTCGACGACACTGCTCT 592
QY 329 ATGCTTTTCATCCGCGCTG 348
Db 593 ACGGATGACCTTGGCGCG 612

RESULT 9

US-09-252-991A-11418
Sequence 11418, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 11418
LENGTH: 1086
TYPE: DNA
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-11418

Query Match 3.6%; Score 39.2; DB 4; Length 1086;
Best Local Similarity 55.0%; Pred. No. 0.019;
Matches 77; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

QY 209 TTTCCGCTGATTTTCGCGGCTCCACCGAATGCTGTCATCGCCCTGTTGTGCGCG 268

Db 222 TGTCCGCCATCGCTTCGCCGCCGCCGCCAGTTGATGCGCATCGCATGCTCAAGGCG 281
QY 269 CAGCGCCCTGAGCGCCGATCGCGCTGACACCATTTGCTGGAATTCGCCACGATTTCT 328
Db 282 GAGCGCGGTTCTTTGATCATCTTCAACACCTCTCTGACCTGCGACGACCTGCTCT 341
QY 329 ATGCGTTTTCATTCGCCGCTG 348
Db 342 ACGGATGAGCTTGGCGCGC 361

RESULT 10
US-09-621-976-8976/c
Sequence 8976, Application US/09621976
Patent No. 6639083
GENERAL INFORMATION:
APPLICANT: Dumas Milne Edwards, J.B.
APPLICANT: Jobert, S.
APPLICANT: Giordano, J.Y.
TITLE OF INVENTION: ESTs and Encoded Human Proteins.
FILE REFERENCE: GENSET 054PR2
CURRENT APPLICATION NUMBER: US/09/621,976
CURRENT FILING DATE: 2000-07-21
NUMBER OF SEQ ID NOS: 19335
SOFTWARE: Patent.pm
SEQ ID NO 8976
LENGTH: 399
TYPE: DNA
ORGANISM: Homo sapiens
US-09-621-976-8976

Query Match 3.6%; Score 39; DB 4; Length 399;
Best Local Similarity 12.6%; Pred. No. 0.012;
Matches 30; Conservative 112; Mismatches 97; Indels 0; Gaps 0;

QY 509 GAGTGGGATCGGACGATTATTCCTTTGAATAAGGCGCTCGAGTTCGCCCTTGTCT 568
Db 242 SAAMRRKMTKMGSRSGSTGYAAMWYKSMCTSRKMYKKRRKRRCTSTKRT 183
QY 569 CTCCTTTGACGCTGACTTGGATTCCTGCGAAGAAAAGACAGATCCCTTCTGTC 628
Db 182 CYGISTYKMKAYTKRRKMTWTYTYYSVMKMTKMTAYTWTBKMTTRTKMT 123
QY 629 TGCTCGGAGTTGAGCTTACCATTTGCTTGTGTAATTCAGGTGAGGCTTTTG 688
Db 122 CTMCWCKTYYMAGTMYRRYRYRYAKRAKMSKRCSTWSTCYCMKYNAKKCWSYMS 63
QY 689 CGGCGCTGCTGATCTTCTGCTGTTGACCATCGGCTACTCTTCTTGGGAAAGCT 747
Db 62 MMKMGSMKMYTYTYTYMMKSMKTYMSMCSYAKCKMTYAKTYTWTCTMTGKRT 4

RESULT 11
US-09-621-976-18033
Sequence 18033, Application US/09621976
Patent No. 6639063
GENERAL INFORMATION:
APPLICANT: Dumas Milne Edwards, J.B.
APPLICANT: Jobert, S.
APPLICANT: Giordano, J.Y.
TITLE OF INVENTION: ESTs and Encoded Human Proteins.
FILE REFERENCE: GENSET 054PR2
CURRENT APPLICATION NUMBER: US/09/621,976
CURRENT FILING DATE: 2000-07-21
NUMBER OF SEQ ID NOS: 19335
SOFTWARE: Patent.pm
SEQ ID NO 18033
LENGTH: 474
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: msec_feature
LOCATION: 16

OTHER INFORMATION: n=a, g, c or t
US-09-621-976-18033

Query Match 3.4%; Score 37; DB 4; Length 474;
Best Local Similarity 14.4%; Pred. No. 0.059;
Matches 49; Conservative 135; Mismatches 156; Indels 0; Gaps 0;

QY 422 CGGCGAGCCCGACGCTGCTCGCGCGGAGCTTATTCGAATGAATAGCTTCACT 481
Db 13 YSGNCTCYCMARKRSTYGRMYKSMRGSSCYKSCMCMKSYCGSYKTTTITWAW 72
QY 482 CCTACTGGATTTGCGCGGTCTACCGAGATGCGATGCGAGATTGCTTTGANA 541
Db 73 WTTTTRKNAARRMSGGKTTTVMGSKKTKSCMGRWKGYYRSWYYCYKACYYW 132
QY 542 TTAAGGCTTCGAGTTGCGCTTGTCTCTCTTTGTCAGGCTTGTGATTCCTGCC 601
Db 133 WKRWYSSCCMYTKGGGSMWTTTMMRRRKYKMTKGRKKTMMMAACYYTWS 192
QY 602 GAAGAAAAGACAGATCCCTTCTCTGCTGCTGCGAGGTTGAGCTTCACCATTTG 661
Db 193 YMMMRRAAARAAKTYCYCMKMTCCMACCCMCMRBRASCMSRBYTMMCYYY 252
QY 662 TGTAATTCAGTACGCGCCATTTGCGCGCTGCTGATCTTGTGGCTGTGAACA 721
Db 253 MMVKGSMYMMWRGMMKEMMYMKKSMWKSMMWRAMWARTYTYTAWYTYTYKR 312
QY 722 TCCGTACTTCTCTTGGGAAAGGCTGTAATGACACT 761
Db 313 MCTYMRKTTYCMWMTYSRWRSMTARGAWMTYMY 352

RESULT 12
US-09-252-991A-3205/c
Sequence 3205, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 3205
LENGTH: 1140
TYPE: DNA
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-3205

Query Match 3.4%; Score 36.8; DB 4; Length 1140;
Best Local Similarity 58.0%; Pred. No. 0.12;
Matches 65; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

QY 211 TCCGCGCTGATTTTGGCGGGTCCACCGAATGCTGTCAGCGCCCTGTTGGGGCA 270
Db 465 TACTTCTGATCAACTTCGGCGGCGCTCGCGCGGCTGATGCGCTGACTTCGGGCTC 406
QY 271 GCGCCCTGCGCGCATGCGGCTCACCATTTGCTGTGAATTCGCCGACG 322
Db 405 GCGGCGAGCGCGGACCTTCTGTGTACCGCGGCTGTCTCTGTAG 354

RESULT 13
US-09-252-991A-2761
Sequence 2761, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.

பெரிய

ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THER

```

NUMBER OF SEQUENCES: 7310
CORRESPONDENCE ADDRESS:
ADDRESSER: GENOME THERAPEUTICS CORPORATION
STREET: 100 Beaver Street
CITY: Waltham
STATE: Massachusetts
COUNTRY: USA
ZIP: 02354

COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660
COMPUTER: PC
OPERATING SYSTEM: <Unknown>
SOFTWARE: ASCII

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/417,884A
FILING DATE: 17-Apr-2003

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/107,532
FILING DATE: 30-Jun-1998
APPLICATION NUMBER: 60/085,598
FILING DATE: 14 May 1998
APPLICATION NUMBER: 60/051571
FILING DATE: July 2, 1997

ATTORNEY/AGENT INFORMATION:
NAME: Arinella, Pamela Deneké
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-012
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781) 893-5007
TELEFAX: (781) 893-8277

INFORMATION FOR SEQ ID NO: 1266:
SEQUENCE CHARACTERISTICS:
LENGTH: 768 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Enterococcus faecium
FEATURE:
NAME/KEY: misc feature
LOCATION: (B) LOCATION 1...768
SEQUENCE DESCRIPTION: SEQ ID NO: 1266:
US-10-417-884A-1266

Query Match      3.7%; Score 39.8; DB 6; Length 768;
Best Local Similarity 49.3%; Pred. No. 0.0085;
Matches 104; Conservative 0; Mismatches 107; Indels 0; Gaps 0;

```

```

APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
FILE REFERENCE: 38-21(5353)B
CURRENT APPLICATION NUMBER: US/10/767,701
CURRENT FILING DATE: 2004-01-29
NUMBER OF SEQ ID NOS: 63128
SEQ ID NO 5263
LENGTH: 619
TYPE: DNA
ORGANISM: Sorghum bicolor
FEATURE:
OTHER INFORMATION: Clone ID: SORBI-28MAY03-CLUS11974_1
US-10-767-701-5263

Query Match      3.3%; Score 35.4; DB 6; Length 619;
Best Local Similarity 54.1%; Pred. No. 0.21;
Matches 72; Conservative 0; Mismatches 61; Indels 0; Gaps 0;

```

```

226 GCGGAGCTCCACGAAATGCTGTCATGCGCCCTGTTGAGGCGAGCGCCCTGGAGGCC 285
132 GCGGAGCTCCACGAAATGCTGTCATGCGCCCTGTTGAGGCGAGCGCCCTGGAGGCC 191
286 ATCGGAGCTCCACGAAATGCTGTCATGCGCCCTGTTGAGGCGAGCGCCCTGGAGGCC 345
192 ATCGGAGCTCCACGAAATGCTGTCATGCGCCCTGTTGAGGCGAGCGCCCTGGAGGCC 251
346 CTGCATGTGTCA 358
252 GTGGGCTGTCTCA 264

RESULT 4
US-10-767-701-10866
Sequence 10866, Application US/10767701
GENERAL INFORMATION:
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
FILE REFERENCE: 38-21(5353)B
CURRENT APPLICATION NUMBER: US/10/767,701
CURRENT FILING DATE: 2004-01-29
NUMBER OF SEQ ID NOS: 63128
SEQ ID NO 10866
LENGTH: 1034
TYPE: DNA
ORGANISM: Sorghum bicolor
FEATURE:
OTHER INFORMATION: Clone ID: SORBI-28MAY03-CLUS4850_1
US-10-767-701-10866

Query Match      3.2%; Score 34.6; DB 6; Length 1034;
Best Local Similarity 49.7%; Pred. No. 0.52;
Matches 88; Conservative 0; Mismatches 89; Indels 0; Gaps 0;

```

```

241 ATGCTGTATGATGCGCCCTGTTGAGGCGAGCGCCCTGGAGCGAGCGCCCTGGAGGCC 300
151 ATGCTGTATGATGCGCCCTGTTGAGGCGAGCGCCCTGGAGCGAGCGCCCTGGAGGCC 210
301 TTGCTGTATGATGCGCCCTGTTGAGGCGAGCGCCCTGGAGCGAGCGCCCTGGAGGCC 360
211 GTGCTGTATGATGCGCCCTGTTGAGGCGAGCGCCCTGGAGCGAGCGCCCTGGAGGCC 270
361 AACCCATGAGCGCTCTCCAGCGGAGCTTCAACCGTCAACCGGAGCGGCGCGTC 417
271 AAGCCATGAGCGCTCTCCAGCGGAGCTTCAACCGTCAACCGGAGCGGCGCGTC 327

RESULT 5
US-10-100-683-983/c

```

Sequence 983, Application US/10100683
GENERAL INFORMATION:
APPLICANT: Rosen, et al.
TITLE OF INVENTION: Human Secreted Proteins
FILE REFERENCE: PS900
CURRENT APPLICATION NUMBER: US/10/100,683
CURRENT FILING DATE: 2002-03-19
PRIOR APPLICATION NUMBER: US 60/040,162
PRIOR FILING DATE: 1997-03-07
PRIOR APPLICATION NUMBER: US 60/043,576
PRIOR FILING DATE: 1997-04-11
PRIOR APPLICATION NUMBER: US 60/047,601
PRIOR FILING DATE: 1997-05-23
PRIOR APPLICATION NUMBER: US 60/056,845
PRIOR FILING DATE: 1997-08-22
PRIOR APPLICATION NUMBER: US 60/043,580
PRIOR FILING DATE: 1997-04-11
PRIOR APPLICATION NUMBER: US 60/047,599
PRIOR FILING DATE: 1997-05-23
PRIOR APPLICATION NUMBER: US 60/056,664
PRIOR FILING DATE: 1997-08-22
PRIOR APPLICATION NUMBER: US 60/043,314
PRIOR FILING DATE: 1997-04-11
PRIOR APPLICATION NUMBER: US 60/047,632
PRIOR FILING DATE: 1997-05-23
PRIOR APPLICATION NUMBER: US 60/056,892
PRIOR FILING DATE: 1997-08-22
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 13468
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO 983
LENGTH: 3173
TYPE: DNA
ORGANISM: Homo sapiens
US-10-100-683-983

Query Match 3.2%; Score 34.6; DB 6; Length 3173;
Best Local Similarity 75.4%; Pred. No. 0.99; Indels 0; Gaps 0;
Matches 43; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

244 CTGTCATCGCCCTCGTTGTGGCGCAGCGCCCTGGGCGCCATCGCGCTCACACCA 300
Db 2360 CAGCTCCCGCGCCCATGTGTGGCTCAGGCGCTCTGGGGCGCATGAGTCTACCCACA 2304

RESULT 6
US-10-100-683-11511
Sequence 11511, Application US/10100683
GENERAL INFORMATION:
APPLICANT: Rosen, et al.
TITLE OF INVENTION: Human Secreted Proteins
FILE REFERENCE: PS900
CURRENT APPLICATION NUMBER: US/10/100,683
CURRENT FILING DATE: 2002-03-19
PRIOR APPLICATION NUMBER: US 60/040,162
PRIOR FILING DATE: 1997-03-07
PRIOR APPLICATION NUMBER: US 60/043,576
PRIOR FILING DATE: 1997-04-11
PRIOR APPLICATION NUMBER: US 60/047,601
PRIOR FILING DATE: 1997-05-23
PRIOR APPLICATION NUMBER: US 60/056,845
PRIOR FILING DATE: 1997-08-22
PRIOR APPLICATION NUMBER: US 60/043,580
PRIOR FILING DATE: 1997-04-11
PRIOR APPLICATION NUMBER: US 60/047,599
PRIOR FILING DATE: 1997-05-23
PRIOR APPLICATION NUMBER: US 60/056,664
PRIOR FILING DATE: 1997-08-22
PRIOR APPLICATION NUMBER: US 60/043,314
PRIOR FILING DATE: 1997-04-11
PRIOR APPLICATION NUMBER: US 60/047,632
PRIOR FILING DATE: 1997-05-23
PRIOR APPLICATION NUMBER: US 60/056,892

PRIOR FILING DATE: 1997-08-22
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 13468
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO 11511
LENGTH: 7071
TYPE: DNA
ORGANISM: Homo sapiens
US-10-100-683-11511

Query Match 3.2%; Score 34.6; DB 6; Length 7071;
Best Local Similarity 75.4%; Pred. No. 1.6; Indels 0; Gaps 0;
Matches 43; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

244 CTGTCATCGCCCTCGTTGTGGCGCAGCGCCCTGGGCGCCATCGCGCTCACACCA 300
Db 785 CAGCTCCCGCGCCCATGTGTGGCTCAGGCGCTCTGGGGCGCATGAGTCTACCCACA 841

RESULT 7
US-10-767-701-7545
Sequence 7545, Application US/10767701
GENERAL INFORMATION:
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(53535)B
CURRENT APPLICATION NUMBER: US/10/767,701
CURRENT FILING DATE: 2004-01-29
NUMBER OF SEQ ID NOS: 63128
SEQ ID NO 7545
LENGTH: 666
TYPE: DNA
ORGANISM: Sorghum bicolor
FEATURE:
OTHER INFORMATION: Clone ID: SORBI-28MAY03-CLUS92927_1
US-10-767-701-7545

Query Match 3.2%; Score 34.4; DB 6; Length 666;
Best Local Similarity 57.4%; Pred. No. 0.47; Indels 46; Gaps 0;
Matches 62; Conservative 0; Mismatches 46; Indels 46; Gaps 0;

215 GCCGATTTTGGCGGGCTCACCGCAATGCTGTGATCGCCCTGTTGTGGCGCAGCG 274
Db 31 GCCCGACACCAAGGAGAGAGACGCCGCTCATCAACCCGTTCTGTGAGCGGCCCGC 90

275 CCTGGGCGCCATCGCGCTACCAATGCTGTGAACTTCCGCCAG 322
Db 91 CGCTGGCGCTCCCTGGCGTGCAGCGCGGCTGCTGTGCGTGGCGAG 138

RESULT 8
US-10-767-701-744/c
Sequence 744, Application US/10767701
GENERAL INFORMATION:
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(53535)B
CURRENT APPLICATION NUMBER: US/10/767,701
CURRENT FILING DATE: 2004-01-29
NUMBER OF SEQ ID NOS: 63128
SEQ ID NO 744
LENGTH: 661
TYPE: DNA
ORGANISM: Sorghum bicolor
FEATURE:
OTHER INFORMATION: Clone ID: SORBI-28MAY03-CLUS124346_1
US-10-767-701-744

Page 4

Query Match	3.1%;	Score 33.8;	DB 7;	Length 24477;
Best Local Similarity	51.9%;	Pred. No.5.9;		
Matches	60;	Conservative	3;	Mismatches 0; Gaps 0;
QY	242	TGCGGCGATCGCGCCTCGTTGTGGGCGACGCGCCTCGGAGCGGCATCGCGCTCACCACAT	301	
D	17603	TGCTGGATCATCGGCTCCCTGATGCTCGCGCTCCTCGGAGGATTCAGAGGAGGGGYYTTC	17663	

[illegible]

```

; PRIOR APPLICATION NUMBER: 60/266,862
; PRIOR FILING DATE: 2001-02-07
; NUMBER OF SEQ ID NOS: 531
; SOFTWARE: PatentIn Ver. 3.2
; SEQ ID NO 500
; LENGTH: 930
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-804-291A-500

Query Match          3.1%      Score 33.2; DB 5; Length 930;
Basic local similarity 54.0%      Pred. No. 1.4;

```

[illegible]

RESULT 12

US-10-021-698A-716
 ; Sequence 716, Application US/10021698A
 ; GENERAL INFORMATION:
 ; APPLICANT: KEITH, TIM
 ; APPLICANT: LITTLE, RANDALL
 ; APPLICANT: VAN BERDEWEGH, PAUL
 ; APPLICANT: DUPUIS, JOSEPH, RICHARD
 ; APPLICANT: DEL MASTRO, RICHARD
 ; APPLICANT: SIMON, JASON
 ; APPLICANT: ALLEN, KRISTINA
 ; APPLICANT: PANDIT, SUNIL
 ; TITLE OF INVENTION: NUCLEOTIDE AND AMINO ACID SEQUENCES RELATING TO
 ; FILE REFERENCE: 2976-404US1
 ; CURRENT APPLICATION NUMBER: US/10/021,698A
 ; PRIOR FILING DATE: 2001-10-22
 ; PRIOR APPLICATION NUMBER: 60/211,749
 ; NUMBER OF SEQ ID NOS: 6160
 ; SOFTWARE: PatentIn 2.1
 ; SEQ ID NO 716
 ; LENGTH: 173233
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: modified base
 ; LOCATION: (88987)..(89086)
 ; OTHER INFORMATION: a, t, c or g
 ; FEATURE:
 ; NAME/KEY: modified base
 ; LOCATION: (167286)..(167385)
 ; OTHER INFORMATION: a, t, c or g
 US-10-021-698A-716

Query Match 3.1%; Score 33.2; DB 6; Length 173233;
 Best Local Similarity 55.1%; Pred. No. 28;

Matches 65; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

QY 952 AGACGTCACTTGTGTCATTCGCGTGCATTCAGTGGGCGCATCTTCTGGG 1011

DB 84488 AGAATCTAACCTTGACCAAAATGCTTACTTACTTGTGTGTTGTTCTCG 84547

QY 1012 GTGCGACGACCTTGTGAGGTTGGCGTGGCAACATCGTTTGTGGACTGGTAA 1069

DB 84548 GTAACCTCAACCTCTGCTGCTGTGCGCTTGTCACTGCTGATTTGCAGATGAA 84605

RESULT 13

US-60-548-091-19934
 ; Sequence 19934, Application US/60548091
 ; GENERAL INFORMATION:
 ; APPLICANT: CARGILL, Michele et al.
 ; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
 ; FILE REFERENCE: CL001506
 ; CURRENT APPLICATION NUMBER: US/60/548,091
 ; CURRENT FILING DATE: 2004-02-27
 ; NUMBER OF SEQ ID NOS: 24433
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 19934
 ; LENGTH: 201
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-60-548-091-19934

Query Match 3.1%; Score 33; DB 7; Length 201;

Best Local Similarity 52.7%; Pred. No. 0.69;
 Matches 69; Conservative 1; Mismatches 61; Indels 0; Gaps 0;

QY 242 TGTGTGTCATCGCCCTCGTTGTGTGGCGCAGCCCTCTGGCGGCATCGGCTCACCACAT 301

DB 46 TGCTGTGTCATCGGCTCCCTGTGTTGGCTCGCTGCTCCCTGGGATTCAGAGAGGGGTTCT 105

QY 302 TGCTGTGAACCTTCGCCACGATATTCTATGCGTTTCAATCCCGCTGCATGTGTCAAAA 361

DB 106 GCTTCAGAGACACCCCGGACAAATTCATCCATGTGCGTCCCGCTGCTGGATCTACCC 165

QY 362 ACCCATTTGCC 372

DB 166 TGCCCTTGCC 176

RESULT 14

US-60-548-091-19938
 ; Sequence 19938, Application US/60548091
 ; GENERAL INFORMATION:
 ; APPLICANT: CARGILL, Michele et al.
 ; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
 ; FILE REFERENCE: CL001506
 ; CURRENT APPLICATION NUMBER: US/60/548,091
 ; CURRENT FILING DATE: 2004-02-27
 ; NUMBER OF SEQ ID NOS: 24433
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 19938
 ; LENGTH: 201
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-60-548-091-19938

Query Match 3.1%; Score 33; DB 7; Length 201;

Best Local Similarity 52.7%; Pred. No. 0.69;
 Matches 69; Conservative 1; Mismatches 61; Indels 0; Gaps 0;

QY 242 TGCTGTGTCATCGCCCTCGTTGTGGCGCAGCCCTCTGGCGGCATCGGCTCACCACAT 301

DB 49 TGCTGTGTCATCGGCTCCCTGTGTTGGCTCGCTGCTCCCTGGGATTCAGAGAGGGGCTTCT 108

QY 302 TGCTGTGAACCTTCGCCACGATATTCTATGCGTTTCAATCCCGCTGCATGTGTCAAAA 361

DB 109 GCTTCAGAGACACCCCGGACAAATTCATCCATGTGCGTCCCGCTGCTGGATCTTACC 168

QY 362 ACCCATTTGCC 372

DB 169 TGCCCTTGCC 179

RESULT 15

US-10-767-701-9054
 ; Sequence 9054, Application US/10767701
 ; GENERAL INFORMATION:
 ; APPLICANT: Kovalic, David K.
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Cao, Yongwei
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
 ; FILE REFERENCE: 38-21(5335)B
 ; CURRENT APPLICATION NUMBER: US/10/767,701
 ; CURRENT FILING DATE: 2004-01-23
 ; NUMBER OF SEQ ID NOS: 63128
 ; SEQ ID NO 9054
 ; LENGTH: 1028
 ; TYPE: DNA
 ; ORGANISM: Sorghum bicolor
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: SORBI-28MAY03-CLUS73927_1
 US-10-767-701-9054

Query Match 3.0%; Score 32.8; DB 6; Length 1028;

Best Local Similarity 54.0%; Pred. No. 2;
 Matches 67; Conservative 0; Mismatches 57; Indels 0; Gaps 0;

QY 609 AAAGCATATCCCTTCTGTGCTGTGCGAGGTTTACGCTTACCATTTGCTTGTGTAAAT 668

Tue Apr 20 06:47:15 2004

us-10-608-504-1_copy_101_1176.rnpn

Page 6

Db	746	ATGACCGTTGCCCTGCTTGGTGGTGGTCTGTCCTTGTGTCGCACATTGACACCCGATGGT	805
Qy	669	TCGAGGTCAAGGCCCTATTTCGGGCGCTGCTATCTCTTGGAGTCGTGAACATCCGGTA	728
Db	806	GCCCCACATGACCCCACTGCGCGGTGTCTGTGTGTGTGGTGGTGGTGGCCATCGTAA	865
Qy	729	CTTC	732
Db	866	GTCC	869

Search completed: April 17, 2004, 20:12:57
Job time : 126.34 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Comogen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 17, 2004, 06:48:18 ; Search time 3484.32 Seconds

(without alignments)
9221.792 Million cell updates/sec

Title: US-10-608-504-1_COPY_101_1176

Perfect score: 1076
Sequence: 1 ggcgaacaaacgaagatgagat.....ttgactgtgtaatttttc 1076

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
EST:*
1: em_estda:*
2: em_esthm:*
3: em_estin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vit:*
21: em_gss_fun:*
22: em_gss_man:*
23: em_gss_pro:*
24: em_gss_pur:*
25: em_gss_rnd:*
26: em_gss_png:*
27: em_gss_vit:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	54.6	5.1	473	28	AO936633 HS041-905
2	40.8	4.2	1201	13	BX356664 BX356664
3	40.8	3.8	1201	13	BX381961 BX381961
4	40.2	3.7	1201	13	BX376097 BX376097

Result	Score	Query Match	Length	DB ID	Description
1	54.6	5.1	473	28	AO936633 HS041-905
2	40.8	4.2	1201	13	BX356664 BX356664
3	40.8	3.8	1201	13	BX381961 BX381961
4	40.2	3.7	1201	13	BX376097 BX376097

ALIGNMENTS

RESULT 1
LOCUS AO936633 473 bp DNA linear GSS 23-AUG-2000
DEFINITION HS041-905 Human Not1 clones Homo sapiens genomic, genomic survey
ACCESSION AO936633
VERSION AO936633.1 GI:7213011
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 473)
Zabarovskiy,E.R., Gikratullin,R., Podowski,R.M., Zabarovska,V.V.,
Kiel,J., Muravenko,O.V., Kozirev,S., Petrenko,L., Skobelva,N.,
Li,J., Protodopov,A., Kashuba,V., Ernberg,I., Winberg,G. and
Walstedt,C.
Not1 clones in the analysis of the human genome
Nucleic Acids Res. 28 (7), 1635-1639 (2000)
MEDLINE 20175728
PUBMED 10710430
COMMENT Contact: Podowski RM
Center for Genomics Research
Karolinska Institute
17177 Stockholm, Sweden
Tel: +46-8-728-6372
Fax: +46-8-337983

Tue Apr 20 06:47:15 2004

us-10-608-504-1_copy_101_1176.rst

Page 2

Email: Raf.Podowski@cgr.ki.se
Class: NotI site.

FEATURES
source location/Qualifiers

1..473
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone_lib="Human NotI clones"

ORIGIN

Query Match 5.1%; Score 54.6; DB 28; Length 473;
Best Local Similarity 56.2%; Pred. No. 0.0009; Indels 0; Gaps 0;
Matches 99; Conservative 0; Mismatches 77;

QY 193 TGGGACAGCCCACTGTTTCCGAGCTGATTTTCGCGGAGCTCCACGAAATGCTGTCATC 252
DB 283 TGGAGAGGCGAGGGGCTCTCGGACATGCTNCGCGGAGCGCGGAGTGTGGCCATC 342
QY 253 GCGCTGCTGTGGGCGGAGCGCCCTGGGCGCCATGCGCTCACCACATTTGCTGTGAAC 312
DB 343 GGCATGCTTAAAGCGAGCGACCACTGCTGATCTGCTACCACTGCTGACT 402
QY 313 TTCGCGACGATTTATGATGCTTTTCATTCGCGCTGATGTCATGCAAAACCCCAT 368
DB 403 TCGACACACTGCTTACGCGCTTGTCCATGCGCCGCTTGTGACAGGACCCCTT 458

RESULT 2
BX356664 1201 bp mRNA linear EST 05-MAY-2003
LOCUS BX356664 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
DEFINITION clone CSOD1015B03 3-PRIME, mRNA sequence.
ACCESSION BX356664
VERSION BX356664.1 GI:30378083
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
TITLE Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
JOURNAL Full-length cDNA libraries and normalization
COMMENT Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/Invitrogen Corporation 1600
Paradey Avenue Genoscope sequence ID : CSOD1015CA02NP1.

FEATURES
source location/Qualifiers

1..1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSOD1015B03"
/issue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 4.2%; Score 45; DB 13; Length 1201;
Best Local Similarity 12.9%; Pred. No. 0.65;
Matches 47; Conservative 166; Mismatches 188; Indels 0; Gaps 0;

QY 207 GTTTCGAGGCTGATTTTCGAGGCTCCACGAAATGCTGATGCGCTCGTGTGG 266
DB 745 STBTSSSTTBTCTTBTSSSTTBTSSSYBSBSBTSSCCTSSBTSSSTTBTSS 804

QY 267 CGAGCGCCCTGGGCGGCGCTCAGCCATTCGATGCTGTCGACTTCGCCACTATT 326
DB 805 SBTCTSSSSSSSBTTSTTSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTST 864

QY 327 CTATGCGTTTCAATCCCGCTGATGTCGCAAAAACCCCATTTCCGCTTCAATCG 386
DB 865 YSSBSST 924

QY 387 TTTCGCGCTTATCGACGAGCTTACGAGCTGCTGCGGCGGCGGCGGCTGTCGAC 446
DB 925 BTSSSBSTTSSSSSBSTSSSBSTSSSBSTSSSBSTSSSBSTSSSBSTSTTT 984

QY 447 GTGCGCACTATCTCAATGCAATAGCTTCACTCTACTGCTGCTTGGCGGCTCAC 506
DB 985 VTSSBTTST 1044

QY 507 CGAGTGGCGATCGACAGTTGATTCCTTTGAATTAAGGCGCTGAGTTCGCCCTTG 566
DB 1045 SSSSSSSSSSSSSSBTTTSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTST 1104

QY 567 CTC 569
DB 1105 TTS 1107

RESULT 3
BX381961 1201 bp mRNA linear EST 08-MAY-2003
LOCUS BX381961 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
DEFINITION clone CSOD1072YF05 3-PRIME, mRNA sequence.
ACCESSION BX381961
VERSION BX381961.1 GI:30453007
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
TITLE Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
JOURNAL Full-length cDNA libraries and normalization
COMMENT Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/Invitrogen Corporation 1600
Paradey Avenue Genoscope sequence ID : CSOD1072CC03NP1.

FEATURES
source location/Qualifiers

1..1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSOD1072YF05"
/issue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 3.8%; Score 40.8; DB 13; Length 1201;
Best Local Similarity 3.0%; Pred. No. 9.4;
Matches 17; Conservative 191; Mismatches 362; Indels 0; Gaps 0;

QY 297 CACATGCTGTGATGATTCGCGACGATTTATGCGTTTCAATCCGCTGATGCTGT 356
DB 406 CCGWNNNNKKKKKKKKKKKKKKKKKKKKKKKKKKKKNNNNNNNNNNNNNNNN 465
QY 357 CAAAAACCCCATGCGCGCTTCTATGCTTTGCGGCTTATGACGAAAGCTACGAGT 416

RESULT 4					
BX376097					
LOCUS		1201 bp	mRNA	linear	EST 08-MAY-2001
DEFINITION	BX376097 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens				
	CDNA clone CSODC022YM12 5-PRIME, mRNA sequence.				
ACCESSION	BX376097				
VERSION	BX376097.1 GI:30434756				
KEYWORDS	EST.				
SOURCE	Homo sapiens				
ORGANISM	Homo sapiens (human)				
	Homo sapiens				

FEATURES	SOURCE
----------	--------

ORIGIN

[illegible]

REFERENCE AUTHORS	TITLE	JOURNAL	COMMENT
RESULTS 5 LOCUS DEFINITION ACCESSION VERSION KEYWORDS SOURCE ORGANISM			
EST Gallus gallus (chicken) Gallus gallus Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianine; Gallus; 1 (bases 1 to 692)	Herault, F., Le Meuth-Netzinger, V., Desert, C., Retout, E., Plumi, F., Klopp, C. and Donaiere, M. Construction and primary characterization of chicken normalized multi-tissue cDNA libraries Unpublished (2003)	INRA, UMR INRA-ENSAE Genetique Animale 65, rue de Saint-Brieuc, RENNES cedex, 35042, FRANCE Tel.: +33 (0) 2.23.48.54.63 Fax: +33 (0) 2.23.48.54.70 Email: Madeleine.Donaiere@rennes.inra.fr	692 bp mRNA linear EST 28-JUL-2001 BX540413 AGENAE Gallus gallus multi-tissues library (gcab) Gallus gallus cDNA clone gcab0001b.d.10 5'prim. mRNA sequence. BX540413 BX540413.1 GI:33294966
at signanapporte@jouy.inra.fr to obtain the chromatogram of this			

sequence.
Plate: 0001 row: d column: 10
Seq primer: M13.
Location/Qualifiers
1..692
/organism="Gallus gallus"

/mol_type="mRNA"
/db_xref="taxon:9031"
/clone="gca0001b.d.10"
/tissue_type="adipose tissue, granulosa, multi-tissues,
oviduct, small follicle, utero-vaginal gland"
/dev_stage="adult"
/lab_host="DH10B"
/clone_id="AGENAE Gallus gallus multi-tissues library
(gca0)"
/note="Vector: pT773D-pac; Clone distribution: AGENAE
Resource centre, Francois Piumi, Francois Piumi, Inra.fr,
INRA, CEA Radiobiologie et Etude du genome (LREG), Domaine
de Vilvert, 78352, Jouy-en-Josas cedex, FRANCE"

ORIGIN

Query Match 3.7%; Score 39.4; DB 13; Length 692;
Best Local Similarity 62.9%; Pred. No. 17;
Matches 61; Conservative 0; Mismatches 36; Indels 0; Gaps 0;

QY 705 CTTGGGCTGTTGACCATCGGACTTCTTCTGGGAAAGGCTGCTAATGCAACTGAT 764
Db 321 CTTGGGCGAGTGGCCATCTGCTCTGCACTTGCACAGACGACGTGACCGCAT 262
QY 765 TTCTCCGTATCTCTCTGTTGTCGAGATGTCAG 801
Db 261 TTTTGCTAGGTACCTTCTGCTGCGAGCTCTGAG 225

RESULT 6 718 bp mRNA linear EST 25-JAN-2001
LOCUS AJ397466 dktz426 Gallus gallus cDNA clone 217r1, mRNA sequence.
DEFINITION AJ397466
ACCESSION AJ397466
VERSION AJ397466.1 GI:7130433
KEYWORDS EST.
SOURCE Gallus gallus (chicken)
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
1 (bases 1 to 718)
Abdrakhmanov, I., Lodygin, D., Gerolt, P., Arakawa, H., Law, A.,
Plachy, J., Korn, B. and Buerstedde, J.M.
A large database of chicken Bursel ESTs as a resource for the
analysis of vertebrate gene function
Genome Res. 10 (12), 2062-2069 (2000)

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
COMMENT
Contact: Buerstedde JM
Cellular Immunology
Heinrich-Pette-Institute
Martinistr. 52, 20251 Hamburg, Germany
Email: URL: http://genetics.hpi.uni-hamburg.de/dt40est.html.
Location/Qualifiers
1..718
/organism="Gallus gallus"
/mol_type="mRNA"
/strain="CB"
/db_xref="taxon:9031"
/clone="217r1"
/tissue_type="Bursa of Fabricius"
/cell_type="bursal lymphocyte"
/dev_stage="2-3 weeks old"
/clone_id="dktz426"

FEATURES

source

ORIGIN

Query Match

3.7%; Score 39.4; DB 9; Length 718;

Best Local Similarity 62.9%; Pred. No. 18;
Matches 61; Conservative 0; Mismatches 36; Indels 0; Gaps 0;

QY 705 CTTGGGCTGTTGACCATCGGACTTCTTCTGGGAAAGGCTGCTAATGCAACTGAT 764
Db 117 CTTGGGCGAGTGGCCATCTGCTCTGCACTTGCACAGACGACGTGACCGCAT 58
QY 765 TTCTCCGTATCTCTCTGTTGTCGAGATGTCAG 801
Db 57 TTTTGCTAGGTACCTTCTGCTGCGAGCTCTGAG 21

RESULT 7

CC684922 743 bp DNA linear GSS 19-JUN-2003
LOCUS OGUAX55TH ZM 0.7.1.5 KB Zea mays genomic clone ZMMBMA0397014,
DEFINITION genomic survey sequence.
ACCESSION CC684922
VERSION CC684922.1 GI:32089698
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 743)
Whitelaw, C.A., Quackenbush, J., Van Aken, S., Uteback, T.,
Reenick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,
Citek, R.W., Nunez, A., Robbins, D. and Lakey, N.
Consortium for Maize Genomics
Unpublished (2002)
Other_GSS: OGUAX55TV
Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TR
Class: sheared ends.
Location/Qualifiers
1..743
/organism="Zea mays"
/mol_type="Genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBMA0397014"
/clone_id="ZM 0.7.1.5 KB"
/note="Vector: pBCSK-; Site: 1; HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"

FEATURES

source

ORIGIN

Query Match 3.7%; Score 39.4; DB 29; Length 743;
Best Local Similarity 53.6%; Pred. No. 18;
Matches 82; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

QY 431 CCGCAGGCTGTCGGCGTGGCGACTTATCTCAATGCAATAGCGTTCACTCTTACG 490
Db 475 CCGCGCGCTCCGCGCGTTCGCTCCGACTCTCATCTCGCATGCGCAATCTCTAG 534
QY 491 TATTCGCGGCTCTCACCGAGTGGGAGTCCGAGATTCCTTTGAATTAAGGCGC 550
Db 535 GTTTCGCGGCGCGCACCGAGTCCAGCTCGGCTTGGCGAGCGCTCTAAGCTTACAGTC 594
QY 551 TCGAGTTCGCCCTTGTCTCTCTCTTGTGTCAGCC 583
Db 595 ACGAGAGGCGGCTTGTCTCTCTCTCCGCTGCGCC 627

RESULT 8

BU448643 756 bp mRNA linear EST 29-NOV-2002
LOCUS BU448643
DEFINITION 603211326F1 CSORBN13 Gallus gallus cDNA clone Chest191c7 5', mRNA

sequence.
 B048643 GI:25937954
 VERSION B048643.1
 KEYWORDS EST.
 SOURCE Gallus gallus (chicken)
 ORGANISM Gallus gallus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.
 1 (bases 1 to 756)
 REFERENCE Boardman, P.E., Sanz-Ezquerro, J., Overton, I.M., Burt, D.W., Bosch, E., Fong, W.T., Tickle, C., Brown, W.R.A., Wilson, S.A. and Hubbard, S.J. A Comprehensive Collection of Chicken cDNAs
 JOURNAL Curr. Biol. 12 (22), 1965-1969 (2002)
 MEDLINE 2235534
 PUBMED 12445392
 COMMENT Contact: Simon Hubbard
 Department of Biomolecular Sciences
 University of Manchester Institute of Science and Technology (UMIST)
 PO Box 88, Manchester, M60 1QD, UK
 Tel: 01612008930
 Fax: 01612360409
 Email: Simon.Hubbard@umist.ac.uk.
 Location/Qualifiers
 1..756
 /organism="Gallus gallus"
 /mol_type="mRNA"
 /strain="Layer"
 /db_xref="taxon:9031"
 /clone="CBST191C7"
 /sex="Female"
 /dev_stage="adult"
 /lab_host="DH10B"
 /clone_lib="CBSEB0N113"
 /note="Organ: ovary; Vector: pBluescript II KS(+); Site_1: EcoRI; Site_2: NotI; This normalized library was constructed from 1 million independent clones. cDNA synthesis was initiated using an oligo(dT) primer, using methylated C in the first strand synthesis reaction. Following this first strand reaction, double-stranded cDNA was blunt-ended, ligated to NotI adaptors, digested with EcoRI, size-selected, and cloned into the NotI and EcoRI compatible sites of a custom modified MCS of the pBluescript (KS+) vector. The library was normalized in 2 rounds using conditions adapted from Soares et al., PNAS (1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6 (1996): 791, except that a significantly longer reannealing hybridization was used."

ORIGIN
 Query Match 3.7%; Score 39.4; DB 13; Length 756;
 Best Local Similarity 62.9%; Pred. No. 18;
 Matches 61; Conservative 0; Mismatches 36; Indels 0; Gaps 0;

QY 705 CTTGGGTCGTGACCATCGGACTTCTCTGGGAAAGGCTGCTAATGCAACTGAT 764
 |||||
 DB 270 CTTGGGTCGTGACCATCGGACTTCTCTGGGAAAGGCTGCTAATGCAACTGAT 211
 |||||
 QY 765 TTCTCCTGATTTCTCTGTTGTCGAGATGTCAG 801
 |||||
 DB 210 TTTTGCTAAGTGCACCTTCTGCTGACGCTGACG 174
 |||||

RESULT 9
 CNS009WA/c 787 bp DNA linear GSS 03-JUN-1999
 LOCUS Drosophila melanogaster genome survey sequence T7 end of BAC #
 DEFINITION BACR20X04 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.
 ACCESSION AL054013
 VERSION AL054013.1 GI:4935590
 KEYWORDS GSS.

SOURCE Drosophila melanogaster (fruit fly)
 ORGANISM Drosophila melanogaster
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
 1 (bases 1 to 787)
 REFERENCE Genoscope.
 Direct Submission
 Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
 - Web : www.genoscope.cns.fr)
 Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see <http://www.fruitfly.org> The BDGP Drosophila melanogaster BAC library was prepared by Kazuhiro Osoegawa and Aaron Mammoler in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain y2; cn bw sp, the same strain used for the BDGP's pl and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
 Location/Qualifiers
 1..787
 /organism="Drosophila melanogaster"
 /mol_type="genomic DNA"
 /db_xref="taxon:7227"
 /clone="BACR20X04"
 /clone_lib="RPCI-98"
 /note="end : T7"

ORIGIN
 Query Match 3.7%; Score 39.4; DB 29; Length 787;
 Best Local Similarity 45.5%; Pred. No. 18;
 Matches 76; Conservative 15; Mismatches 76; Indels 0; Gaps 0;

QY 552 CGAGTGGCCCTTGTCTCTCTCTGTCACGCTGATTTGATTCGCCGAGAAA 611
 |||||
 DB 334 CTATTTGCTCTTCTCTCTCTTGTGTAAMKMGMMCTGGAATCCCTKXAGARAAAC 275
 |||||
 QY 612 GCAATCCCTTCTCTGTCGTCGAGGTTGACCTTACCATGCTCTTGTAATCC 671
 |||||
 DB 274 AYTGTGTTKTTTCYACWGRCTTAAGGTAATTTCTTAATTTTATTTTAA 215
 |||||
 QY 672 AGTCAGCCCTATTTCGGCGCTGATCTTCTGAGTCTGTA 718
 |||||
 DB 214 TAATTTGCCATATGCTTATATGATTTTATTTATGTCGGCTGA 168
 |||||

RESULT 10
 Bx425603 885 bp mRNA linear EST 15-MAY-2003
 LOCUS Bx425603 Homo sapiens NEUROBLASTOMA Homo sapiens cDNA clone
 DEFINITION C10B0222A07 3-PRIME, mRNA sequence.
 ACCESSION Bx425603
 VERSION Bx425603.1 GI:30770486
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 1 (bases 1 to 885)
 REFERENCE Li, W.B., Gruber, C., Jessee, J. and Polayes, D. Full-length cDNA libraries and normalization unpublished (2001)
 COMMENT Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 EVRY cedex - France
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr

P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES

source

1..925
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACR19D16"
/clone_1b="RPCI-98"
/note="end : TET3"

ORIGIN

Query Match 3.6%; Score 39; DB 29; Length 925;
Best Local Similarity 14.2%; Pred. No. 26;
Matches 37; Conservative 117; Mismatches 106; Indels 0; Gaps 0;

195 GGCAGCCCGCATGTTTCGCGCTGATTTTCGCGGCTCCACCAATGCTGTCATGAC 254
:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
638 SSSSSSTSSSTSSSTSSKSSSSSSSSSSSYTTKSTASGSGWSAGSGSGTSGTSS 697
:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
255 CCTCGTTGGGCGAGCGCCCTGGGCGCATCGCGCTCACCACTTGTGTAATT 314
:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
698 SSSSSSTSSSSSVSSGSKSSTBSGSSSSSSSTSSBBSCTSTSSSSSSSYSSGT 757
:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
315 CGCGCAGATTCATGCGTTTTCATTCGCGGTGATGTCGCAAAACCCCATTTGCCG 374
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
758 GCGCTCCGCSYSSSTSSSSSTWSGTSSSSSSVGTSSSDTSTCCSCCTTCCTST 817
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
375 TTCTATTGCGTTTCGCGCTTATGACGACGACCTACGACGACTGCTGCGCGCCGCGC 434
:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
818 YAMBCTSTSCGSSSSSGKGVTKCGCGCSTTNGMBTSSACSSSSSSSSSVSS 877
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
435 AGGCTGTGCGCGTGCAC 454
:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
878 SSKSSASSSSSVSSGSSSVS 897
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

RESULT 13

BX361080

LOCUS

BX361080 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Genoscope
Contact: Genoscope
Genoscope Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: segre@genoscope.cns.fr; Web: www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of Invitrogen. Contact: Feng Liang Email: filang@lifetech.com URL: <http://fulllength.invitrogen.com/> Invitrogen Corporation 1600 Faraday Avenue Genoscope sequence ID : CS0D1079D608NP1.
Location/Qualifiers
1..1201
/organism="Homo sapiens"
/mol_type="RNA"
/db_xref="taxon:9606"
/clone="CS0D1079YNI6"
/issue_type="PLACENTA COT 25-NORMALIZED"
/clone_1b="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was placed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was

ORIGIN

Query Match 3.6%; Score 38.6; DB 13; Length 1201;
Best Local Similarity 10.5%; Pred. No. 38;
Matches 34; Conservative 147; Mismatches 142; Indels 0; Gaps 0;

digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

19 ATTCATTCAAGCTGAGAGTGTCGCATCCAGAGCAGCCTGAGAACGATGATAAGT 78
:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
825 MDDMMMMMMKKNDKKKKNDKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKK 884
:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
79 TATCGCGCTACGAAATCGCGCAAGCTCTAAACCTCTGCTGAGGTTGGCATG 138
:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
885 KKKKKDAAKK 944
:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
139 TACCGATGATGATGCGTTGCTCTGTTGTTATTCATATACGAGCTACGAGTGGCA 198
:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
945 KKKKKKKKKKKTT 1004
:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
199 GCCCGCATGTTTCCGCGCGATTTTCCGCGGCTCCACCGAATGCTGTCATGCGCCTC 258
:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
1005 KNDMMMMMMKK 1064
:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
259 GTTGTGGCGCAGCGCCCTGGGCGCATCGCTACACCATGCTGTGTAATTCCGC 318
:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
1065 KKKKKKKCCCBKCCGCCCKKCKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKK 1124
:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
319 CAGCATTTCTATGCGTTTCATT 341
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
1125 CCKKKKKCKKKKKKKKKKKKK 1147
|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

RESULT 14

CB657749

LOCUS

CB657749 OSJNEC13F02.f OSJNEC Oryza sativa (japonica cultivar-group) cDNA

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Oryza sativa (japonica cultivar-group)
Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.
1 (bases 1 to 835)
Jantasuriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,E.,
Kudrna,D., Dean,R., Soderlund,C., Wang,R. and Wang,G.
Large-scale identification of ESTs involved in the interaction
between rice and Magnaporthe grisea
Unpublished (2003)
Contact: Rod Wang
Arizona Genomics Institute
University of Arizona
Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
85721-0088, USA
Tel: 520 626 3967
Fax: 520 621 9288
Email: <http://genome.arizona.edu>
PCR Primers
FORWARD: gga aac cga cgg cca gtc
BACKWARD: gga aac agc tat gac cat g
Plate: 13 row: F column: 02
Seq primer: gta aac cga cgg cca gtcg.
Location/Qualifiers
1..835
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="RNA"
/cullivar="Nipponbare"
/db_xref="taxon:39947"
/clone="OSJNEC13F02"
/issue_type="leaf"

Tue Apr 20 06:47:15 2004

us-10-608-504-1_copy_101_1176.rst

Page 8

ORIGIN

```
/dev_stage="3 week"
/lab_host="DH10B"
/clone_lib="OSJunc"
/notes="Vector: pBluescript II KS +; Site 1: EcoRI; Site 2:
XhoI; 6 hrs after inoculation with Rice Blast (C9240-1)"
```

Query Match 3.5%; Score 38; DB 14; Length 835;
Best Local Similarity 51.1%; Pred. No. 46;
Matches 89; Conservative 0; Mismatches 85; Indels 0; Gaps 0;

QY 252 CCGCCCTGCTTGGGCGCAGCCCGCCCTGAGCGCTCACCACATTGCTGTGA 311
DB 79 CACACCGCGCTCGGCGCGCGCGCTGCTGCGCGCGCGCGCGCGCGCGCTC 138
QY 312 CTTCCGCGCAGCTATTGCTTATCGCTTTCATCCCGCTGCTGCTGCTGCTG 371
DB 139 ATGCGCGCGCGCGCTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCTGCT 198
QY 372 CCGTTTCTATTGCGCTTTCGCGCTTATCGAGAGCTTACGAGTCACTGCGGC 425
DB 199 CTCCGCTTACGCGCGCTTCTGCGCGAGGACGCTGCGCTGCGCGCTGCGGC 252

RESULT 15
CNS005NG/c 995 bp DNA linear GSS 03-JUN-1999
LOCUS Drosophila melanogaster genome survey sequence TET3 end of BAC #
DEFINITION BACR12604 of RPCL-98 library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
AL060428
GI:4943359
GSS.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Drosophila melanogaster (fruit fly)
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 995)
Genoscope.
Direct Submission
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)

REFERENCE
AUTHORS
TITLE
JOURNAL
Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see <http://www.fruitfly.org> The BDGP Drosophila
melanogaster BAC library was prepared by Kazutoyo Oosawa and
Aaron Mammoxer in Pletier de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPCL-98 and was constructed by partial
EcoRI digestion of Drosophila DNA provided by the BDGP from the
isogenic strain Y2; cn bw sp, the same strain used for the BDGP's
P1 and EST libraries. A more detailed description of the library
and how to order individual BAC clones, the entire library, or
filters for hybridization from the BACPAC Resource Center can be
found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

COMMENT

FEATURES
Source
1. .995
Location/Qualifiers
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACR12604"
/clone_lib="RPCL-98"
/note="end : TET3"

ORIGIN

Query Match 3.5%; Score 37.8; DB 29; Length 995;
Best Local Similarity 34.0%; Pred. No. 57;
Matches 54; Conservative 41; Mismatches 64; Indels 0; Gaps 0;

QY 916 CCGCATCAAGCTTTCGACGAATGCGATCTGTAAGACTTAACCTTGTCTCATTTG 975
DB 872 CAGTMTTACSGAGCGCTGTGTBGAAGAASASATTTGVRSTSGCTTTTMTTAAAT 813
QY 976 CCGTTGCGATTACAGTGTGTGGCGCATCTTCTTGGCGGTGACGACCTTTGAGCGTTG 1035
DB 812 CSKTTSSBTWGAACSCRMTTTSGSWASWTGCTGKKKGAGSAGRGAAAGGCGGAKKSYTB 753
QY 1036 GCGCTGCAACCATCGTTTGTGTGACAGTGAATCTT 1074
DB 752 SSGBTGTGKXTSARKKTTTGCTGGGTGGDNTTCTTT 714

Search completed: April 17, 2004, 20:01:41
Job time : 3485.32 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 19, 2004, 10:57:18 ; Search time 3191.51 Seconds

(without alignments)
10226.291 Million cell updates/sec

Title: US-10-608-504-2

Perfect score: 753

Sequence: 1 gtgcacaaacgaagat.....tcttg99aaagctgctaa 753

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 3470272 segs, 2167151695 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : GenEmbl.*
1: gb pa.*
2: gb hcg.*
3: gb in.*
4: gb om.*
5: gb ov.*
6: gb pa.*
7: gb pl.*
8: gb pr.*
9: gb ro.*
10: gb st.*
11: gb st.*
12: gb st.*
13: gb st.*
14: gb st.*
15: gb st.*
16: gb st.*
17: gb st.*
18: gb st.*
19: gb st.*
20: gb st.*
21: gb st.*
22: gb st.*
23: gb st.*
24: gb st.*
25: gb st.*
26: gb st.*
27: gb st.*
28: gb st.*
29: gb st.*
30: gb st.*
31: gb st.*
32: gb st.*
33: gb st.*
34: gb st.*
35: gb st.*
36: gb st.*
37: gb st.*
38: gb st.*
39: gb st.*
40: gb st.*
41: gb st.*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	753	100.0	753	AR391954	AR391954 Sequence
2	753	100.0	753	AX120373	AX120373 Sequence
3	753	100.0	753	AX137710	AX137710 Sequence
4	753	100.0	753	BD014992	BD014992 Nucleotide
5	753	100.0	753	BD162490	BD162490 Novel pol
6	753	100.0	1271	AR391953	AR391953 Sequence
7	753	100.0	1271	AR391956	AR391956 Sequence
8	753	100.0	1271	AX137709	AX137709 Sequence
9	753	100.0	1271	AX137714	AX137714 Sequence
10	753	100.0	1271	BD014991	BD014991 Nucleotide
11	753	100.0	1271	BD014994	BD014994 Nucleotide
12	753	100.0	2105	AR454053	AR454053 Coryneb
13	753	100.0	340000	AP005274	AP005274 Coryneb
14	753	100.0	349980	AX120085	AX120085 Sequence
15	273.6	36.3	397	AX137083	AX137083 Sequence
16	273.6	36.3	778	AX137085	AX137085 Sequence
17	149.6	19.9	87340	AP005224	AP005224 Coryneb
18	132.6	17.6	53793	AX123492	AX123492 Sequence
19	127.4	16.9	711	BD165609	BD165609 Novel pol
20	127.4	16.9	309400	AX127153	AX127153 Sequence
21	127.4	16.9	325651	AP005283	AP005283 Coryneb
22	127.4	16.9	715	AX137077	AX137077 Sequence
23	85	11.3	715	BD013995	BD013995 Novel nuc
24	85	11.3	715	BD013995	BD013995 Novel nuc
25	59.8	7.8	349926	AX137083	AX137083 Sequence
26	58.6	7.6	9888	AX137083	AX137083 Sequence
27	57.2	7.2	9888	AX137083	AX137083 Sequence
28	54.4	6.9	11025	AX137083	AX137083 Sequence
29	51.8	6.5	11025	AX137083	AX137083 Sequence
30	51.8	6.5	11025	AX137083	AX137083 Sequence
31	48	6.2	2000	AX137083	AX137083 Sequence
32	47	6.2	2000	AX137083	AX137083 Sequence
33	46.6	6.2	10733	AX137083	AX137083 Sequence
34	46.6	6.2	10733	AX137083	AX137083 Sequence
35	46.6	6.2	10733	AX137083	AX137083 Sequence
36	46.2	6.1	7218	AX137083	AX137083 Sequence
37	45	6.1	215050	AX137083	AX137083 Sequence
38	45.8	6.1	20342	AX137083	AX137083 Sequence
39	45.4	6.0	348624	AX137083	AX137083 Sequence
40	45.4	6.0	348624	AX137083	AX137083 Sequence
41	44.6	5.9	329709	AX137083	AX137083 Sequence
42	44.2	5.9	208524	AX137083	AX137083 Sequence
43	44.2	5.9	264050	AX137083	AX137083 Sequence
44	44.2	5.9	300592	AX137083	AX137083 Sequence
45	43.4	5.8	346294	AX137083	AX137083 Sequence

ALIGNMENTS

RESULT 1
LOCUS AR391954
DEFINITION Sequence 2 from patent US 6613545.
ACCESSION AR391954
VERSION AR391954.1 GI:40115725
KEYWORDS
SOURCE
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 753)
AUTHORS Kemmerknecht, N., Sahm, H., Eggeling, L. and Pfeifferle, N.
TITLE Nucleotide sequences coding for the export of branched chain amino acids, process for the isolation thereof and use thereof
JOURNAL Patent: US 6613545-A 2 02-SEP-2003;

FEATURES

source

Location/Qualifiers
1..753
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN

Query Match 100.0%; Score 753; DB 6; Length 753;
Best Local Similarity 100.0%; Pred. No. 9.8e-185;
Matches 753; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 GTGCAAAAAACGCAAGATTCATTCAAGCTGAGGTGTCGCTCAAGGACCCCTG 60
1 GTGCAAAAAACGCAAGATTCATTCAAGCTGAGGTGTCGCTCAAGGACCCCTG 60
61 GAACCAATGATTAAGTTATCGGCGCTAGCAAAATGCGGCAAGGTCTAAAACTCCCT 120
61 GAACCAATGATTAAGTTATCGGCGCTAGCAAAATGCGGCAAGGTCTAAAACTCCCT 120
61 GAACCAATGATTAAGTTATCGGCGCTAGCAAAATGCGGCAAGGTCTAAAACTCCCT 120
121 GCTGCAAGTTGGGCAATGTAACCGATTGATTCGTTGCTCTTGTATTCAATAC 180
121 GCTGCAAGTTGGGCAATGTAACCGATTGATTCGTTGCTCTTGTATTCAATAC 180
181 GGTACGATGATGTTGGGCAAGCCCACTGTTTCCGCGCTAATTTCCGCGGCTCCACCGA 240
181 GGTACGATGATGTTGGGCAAGCCCACTGTTTCCGCGCTAATTTCCGCGGCTCCACCGA 240
181 GGTACGATGATGTTGGGCAAGCCCACTGTTTCCGCGCTAATTTCCGCGGCTCCACCGA 240
241 ATGCTGATCATCGCCCTCGTTGTTGGGCGACGCGCCCTGCGGCGCATCGCGCTACACA 300
241 ATGCTGATCATCGCCCTCGTTGTTGGGCGACGCGCCCTGCGGCGCATCGCGCTACACA 300
301 TTGCTGATGATTCGCGCACTGATTCATGCTTTTCAATCCGCTGCAATGTTGCAAA 360
301 TTGCTGATGATTCGCGCACTGATTCATGCTTTTCAATCCGCTGCAATGTTGCAAA 360
361 AACCCATTGCGCGCTTCTATTCGATTTGCGGCTTATCGACGAAGCTACGACGACT 420
361 AACCCATTGCGCGCTTCTATTCGATTTGCGGCTTATCGACGAAGCTACGACGACT 420
421 GCGGCCAGGCCCGGAGCTGTTGCGGCGTGGCGACTTATCTCAATGCAAAATAGCGTTTCA 480
421 GCGGCCAGGCCCGGAGCTGTTGCGGCGTGGCGACTTATCTCAATGCAAAATAGCGTTTCA 480
481 TCCCTACGAGTATTCGCGGCTGCTCAACGAGTGGCGCATGCGAGTTGATTCCTTTGAA 540
481 TCCCTACGAGTATTCGCGGCTGCTCAACGAGTGGCGCATGCGAGTTGATTCCTTTGAA 540
481 TCCCTACGAGTATTCGCGGCTGCTCAACGAGTGGCGCATGCGAGTTGATTCCTTTGAA 540
541 ATTAAAGGCTCGAGTTCGCGCTTGTCTCTCTTTGTCAAGCTGACTTTGATTCCTG 600
541 ATTAAAGGCTCGAGTTCGCGCTTGTCTCTCTTTGTCAAGCTGACTTTGATTCCTG 600
601 CGAAGCAAAAAGCAGATTCCTTCTCTGCTCTGCGAGTTGAGCTTCAACATTCCTT 660
601 CGAAGCAAAAAGCAGATTCCTTCTCTGCTCTGCGAGTTGAGCTTCAACATTCCTT 660
661 GTGTAATTCAGGTCAGGCTTATTTGCGGCGCTGCTGATTCCTTGGGCTGTGAC 720
661 GTGTAATTCAGGTCAGGCTTATTTGCGGCGCTGCTGATTCCTTGGGCTGTGAC 720
721 ATCCGTAATTCCTTCTTGGGCAAAAGCTGCTAAA 753
721 ATCCGTAATTCCTTCTTGGGCAAAAGCTGCTAAA 753

RESULT 2

AX120373 753 bp DNA linear PAT 11-MAY-2001
LOCUS AX120373
DEFINITION Sequence 289 from Patent EP108790.
ACCESSION AX120373
VERSION AX120373.1 GI:14037088
KEYWORDS
SOURCE Corynebacterium glutamicum
Corynebacterium glutamicum
ORGANISM Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;

REFERENCE

Corynebacterineae; Corynebacteriaceae; Corynebacterium.
1 Nakagawa, S., Mizoguchi, H., Ando, S., Hayaishi, M., Ochiai, K.,
Tokoi, H., Tateishi, N., Senoh, A., Ikeda, M. and Ozaki, A.,
Novel polynucleotides
Patent: EP 1108790-A 289 20-JUN-2001;
KYOWA HAKKO KOGYO CO., LTD. (JP)

FEATURES

source

Location/Qualifiers
1..753
/organism="Corynebacterium glutamicum"
/mol_type="unassigned DNA"
/db_xref="taxon:1718"

ORIGIN

Query Match 100.0%; Score 753; DB 6; Length 753;
Best Local Similarity 100.0%; Pred. No. 9.8e-185;
Matches 753; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 GTGCAAAAAACGCAAGATTCATTCAAGCTGAGGTGTCGCTCAAGGACCCCTG 60
1 GTGCAAAAAACGCAAGATTCATTCAAGCTGAGGTGTCGCTCAAGGACCCCTG 60
61 GAACCAATGATTAAGTTATCGGCGCTAGCAAAATGCGGCAAGGTCTAAAACTCCCT 120
61 GAACCAATGATTAAGTTATCGGCGCTAGCAAAATGCGGCAAGGTCTAAAACTCCCT 120
61 GAACCAATGATTAAGTTATCGGCGCTAGCAAAATGCGGCAAGGTCTAAAACTCCCT 120
121 GCTGCAAGTTGGGCAATGTAACCGATTGATTCGTTGCTCTTGTATTCAATAC 180
121 GCTGCAAGTTGGGCAATGTAACCGATTGATTCGTTGCTCTTGTATTCAATAC 180
181 GGTACGATGATGTTGGGCAAGCCCACTGTTTCCGCGCTAATTTCCGCGGCTCCACCGA 240
181 GGTACGATGATGTTGGGCAAGCCCACTGTTTCCGCGCTAATTTCCGCGGCTCCACCGA 240
181 GGTACGATGATGTTGGGCAAGCCCACTGTTTCCGCGCTAATTTCCGCGGCTCCACCGA 240
241 ATGCTGATCATCGCCCTCGTTGTTGGGCGACGCGCCCTGCGGCGCATCGCGCTACACA 300
241 ATGCTGATCATCGCCCTCGTTGTTGGGCGACGCGCCCTGCGGCGCATCGCGCTACACA 300
301 TTGCTGATGATTCGCGCACTGATTCATGCTTTTCAATCCGCTGCAATGTTGCAAA 360
301 TTGCTGATGATTCGCGCACTGATTCATGCTTTTCAATCCGCTGCAATGTTGCAAA 360
361 AACCCATTGCGCGCTTCTATTCGATTTGCGGCTTATCGACGAAGCTACGACGACT 420
361 AACCCATTGCGCGCTTCTATTCGATTTGCGGCTTATCGACGAAGCTACGACGACT 420
421 GCGGCCAGGCCCGGAGCTGTTGCGGCGTGGCGACTTATCTCAATGCAAAATAGCGTTTCA 480
421 GCGGCCAGGCCCGGAGCTGTTGCGGCGTGGCGACTTATCTCAATGCAAAATAGCGTTTCA 480
481 TCCCTACGAGTATTCGCGGCTGCTCAACGAGTGGCGCATGCGAGTTGATTCCTTTGAA 540
481 TCCCTACGAGTATTCGCGGCTGCTCAACGAGTGGCGCATGCGAGTTGATTCCTTTGAA 540
481 TCCCTACGAGTATTCGCGGCTGCTCAACGAGTGGCGCATGCGAGTTGATTCCTTTGAA 540
541 ATTAAAGGCTCGAGTTCGCGCTTGTCTCTCTTTGTCAAGCTGACTTTGATTCCTG 600
541 ATTAAAGGCTCGAGTTCGCGCTTGTCTCTCTTTGTCAAGCTGACTTTGATTCCTG 600
601 CGAAGCAAAAAGCAGATTCCTTCTCTGCTCTGCGAGTTGAGCTTCAACATTCCTT 660
601 CGAAGCAAAAAGCAGATTCCTTCTCTGCTCTGCGAGTTGAGCTTCAACATTCCTT 660
661 GTGTAATTCAGGTCAGGCTTATTTGCGGCGCTGCTGATTCCTTGGGCTGTGAC 720
661 GTGTAATTCAGGTCAGGCTTATTTGCGGCGCTGCTGATTCCTTGGGCTGTGAC 720
721 ATCCGTAATTCCTTCTTGGGCAAAAGCTGCTAAA 753
721 ATCCGTAATTCCTTCTTGGGCAAAAGCTGCTAAA 753

RESULT 3

AX137710

LOCUS	AX137710	753 bp	DNA	linear	PAT 30-MAY-2000
DEFINITION	Sequence 2 from Patent EP1096010.				
ACCESSION	AX137710				
VERSION	AX137710.1	GI:14273889			
KEYWORDS					
SOURCE					
ORGANISM	Corynebacterium glutamicum				
	Corynebacterium glutamicum				
	Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;				
	Corynebacteriaceae; Corynebacteriaceae; Corynebacterium.				
REFERENCE	1				
AUTHORS	Kernerkecht, N., Eggeling, L., Salm, H. and Pfeifferle, W.				
TITLE	Nucleotide sequences coding for branched-chain amino acids export proteins, method for isolating them and their use				
JOURNAL	Patent: EP 1096010-A 2 02-MAY-2001;				
DEPOSITOR	Degussa AG (DE) ; FORSCHUNGSZENTRUM JUELICH GMBH (DE)				
FEATURES	Location/Qualifiers				
source	1..753				
	/organism="Corynebacterium glutamicum"				
	/mol_type="unassigned DNA"				
	/db_xref="taxon:1718"				
	/note="ATCC14752"				
	1..>753				
	/note="unnamed protein product; brnF"				
	/codon_start=1				
	/transl_table=11				
	/protein_id="CAC39925.1"				
	/db_xref="GI:14273890"				
	/db_xref="REMBL:CAC39925"				
	/translation="MOCKOEHSLEVSFSKALPEPDDGYRYEIAOGLKTSLAGI GMVPIGIAFGILVIOGYEWMAAPLPSGILGASTEMIVILVGAAPLALITLLT VAFRVPVAFSPFPLHVNKSNIPARFVSVPPLIDEAYVAATAPAGSAMRLISMOIAFPI SIWVGAGTGAVALIPEIKIGEPALOSLPVTLTDSCTKQIPSLILAGLSFTT ALVIVPGALFPAALIFLGLTRIFLGRKAK"				
ORIGIN					
Query Match	100.0%; Score 753; DB 6; Length 753;				
Best Local Similarity	100.0%; Pred. No. 9,8e-185;				
Matches	753; Conservative 0; Mismatches 0; Indels 0; Gaps 0				
1	GTGCAAAAAAGCAGAGATTCAATTCAGCCTTGAGAGTGTGGCATTCAGGACGCTTG	60			
2	GTGCAAAAAAGCAGAGATTCAATTCAGCCTTGAGAGTGTGGCATTCAGGACGCTTG	60			
3	GTGCAAAAAAGCAGAGATTCAATTCAGCCTTGAGAGTGTGGCATTCAGGACGCTTG	60			
4	GTGCAAAAAAGCAGAGATTCAATTCAGCCTTGAGAGTGTGGCATTCAGGACGCTTG	60			
5	GTGCAAAAAAGCAGAGATTCAATTCAGCCTTGAGAGTGTGGCATTCAGGACGCTTG	60			
6	GTGCAAAAAAGCAGAGATTCAATTCAGCCTTGAGAGTGTGGCATTCAGGACGCTTG	60			
7	GTGCAAAAAAGCAGAGATTCAATTCAGCCTTGAGAGTGTGGCATTCAGGACGCTTG	60			
8	GTGCAAAAAAGCAGAGATTCAATTCAGCCTTGAGAGTGTGGCATTCAGGACGCTTG	60			
9	GTGCAAAAAAGCAGAGATTCAATTCAGCCTTGAGAGTGTGGCATTCAGGACGCTTG	60			
10	GTGCAAAAAAGCAGAGATTCAATTCAGCCTTGAGAGTGTGGCATTCAGGACGCTTG	60			
11	GTGCAAAAAAGCAGAGATTCAATTCAGCCTTGAGAGTGTGGCATTCAGGACGCTTG	60			
12	GTGCAAAAAAGCAGAGATTCAATTCAGCCTTGAGAGTGTGGCATTCAGGACGCTTG	60			
13	GTGCAAAAAAGCAGAGATTCAATTCAGCCTTGAGAGTGTGGCATTCAGGACGCTTG	60			
14	GTGCAAAAAAGCAGAGATTCAATTCAGCCTTGAGAGTGTGGCATTCAGGACGCTTG	60			
15	GTGCAAAAAAGCAGAGATTCAATTCAGCCTTGAGAGTGTGGCATTCAGGACGCTTG	60			
16	GTGCAAAAAAGCAGAGATTCAATTCAGCCTTGAGAGTGTGGCATTCAGGACGCTTG	60			
17	GTGCAAAAAAGCAGAGATTCAATTCAGCCTTGAGAGTGTGGCATTCAGGACGCTTG	60			
18	GTGCAAAAAAGCAGAGATTCAATTCAGCCTTGAGAGTGTGGCATTCAGGACGCTTG	60			
19	GTGCAAAAAAGCAGAGATTCAATTCAGCCTTGAGAGTGTGGCATTCAGGACGCTTG	60			
20	GTGCAAAAAAGCAGAGATTCAATTCAGCCTTGAGAGTGTGGCATTCAGGACGCTTG	60			
21	GTGCAAAAAAGCAGAGATTCAATTCAGCCTTGAGAGTGTGGCATTCAGGACGCTTG	60			
22	GTGCAAAAAAGCAGAGATTCAATTCAGCCTTGAGAGTGTGGCATTCAGGACGCTTG	60			
23	GTGCAAAAAAGCAGAGATTCAATTCAGCCTTGAGAGTGTGGCATTCAGGACGCTTG	60			
24	GTGCAAAAAAGCAGAGATTCAATTCAGCCTTGAGAGTGTGGCATTCAGGACGCTTG	60			
25	GTGCAAAAAAGCAGAGATTCAATTCAGCCTTGAGAGTGTGGCATTCAGGACGCTTG	60			
26	GTGCAAAAAAGCAGAGATTCAATTCAGCCTTGAGAGTGTGGCATTCAGGACGCTTG	60			
27	GTGCAAAAAAGCAGAGATTCAATTCAGCCTTGAGAGTGTGGCATTCAGGACGCTTG	60			
28	GTGCAAAAAAGCAGAGATTCAATTCAGCCTTGAGAGTGTGGCATTCAGGACGCTTG	60			
29	GTGCAAAAAAGCAGAGATTCAATTCAGCCTTGAGAGTGTGGCATTCAGGACGCTTG	60			
30	GTGCAAAAAAGCAGAGATTCAATTCAGCCTTGAGAGTGTGGCATTCAGGACGCTTG	60			
31	GTGCAAAAAAGCAGAGATTCAATTCAGCCTTGAGAGTGTGGCATTCAGGACGCTTG	60			
32	GTGCAAAAAAGCAGAGATTCAATTCAGCCTTGAGAGTGTGGCATTCAGGACGCTTG	60			
33	GTGCAAAAAAGCAGAGATTCAATTCAGCCTTGAGAGTGTGGCATTCAGGACGCTTG	60			
34	GTGCAAAAAAGCAGAGATTCAATTCAGCCTTGAGAGTGTGGCATTCAGGACGCTTG	60			
35	GTGCAAAAAAGCAGAGATTCAATTCAGCCTTGAGAGTGTGGCATTCAGGACGCTTG	60			
36	GTGCAAAAAAGCAGAGATTCAATTCAGCCTTGAGAGTGTGGCATTCAGGACGCTTG	60			
37	GTGCAAAAAAGCAGAGATTCAATTCAGCCTTGAGAGTGTGGCATTCAGGACGCTTG	60			
38	GTGCAAAAAAGCAGAGATTCAATTCAGCCTTGAGAGTGTGGCATTCAGGACGCTTG	60			
39	GTGCAAAAAAGCAGAGATTCAATTCAGCCTTGAGAGTGTGGCATTCAGGACGCTTG	60			
40	GTGCAAAAAAGCAGAGATTCAATTCAGCCTTGAGAGTGTGGCATTCAGGACGCTTG	60			
41	GTGCAAAAAAGCAGAGATTCAATTCAGCCTTGAGAGTGTGGCATTCAGGACGCTTG	60			
42	GTGCAAAAAAGCAGAGATTCAATTCAGCCTTGAGAGTGTGGCATTCAGGACGCTTG	60			
43	GTGCAAAAAAGCAGAGATTCAATTCAGCCTTGAGAGTGTGGCATTCAGGACGCTTG	60			

Db	481	TCCTACTGGGTAATTCGACGGCTTCACCGGAGTGGCATCGACAGATGTGATTCCTTTGAA	540
Qy	541	ATTAAAGGCTCGAATTCGCTTTGCTCTCTTTTGACAGCTGACTTGATTCCTG	600
Db	541	ATTAAAGGCTCGAATTCGCTTTGCTCTCTTTTGACAGCTGACTTGATTCCTG	600
Qy	601	CGAACGAAAAAGAGATCCCTTCTCTGCTGCTCGACAGTTGAGCTTCACCATTCCTT	660
Db	601	CGAACGAAAAAGAGATCCCTTCTCTGCTGCTCGACAGTTGAGCTTCACCATTCCTT	660
Qy	661	GTGGTAATTCAGATCAGGCGCTTAATTTGGCGCGCTGTGTAATCTTTGGGTCTTGACC	720
Db	661	GTGGTAATTCAGATCAGGCGCTTAATTTGGCGCGCTGTGTAATCTTTGGGTCTTGACC	720
Qy	721	ATCCGGACTCTCTCTTGGGAAAAGCTGCTAAA	753
Db	721	ATCCGGACTCTCTCTTGGGAAAAGCTGCTAAA	753
RESULT 4			
BD014992		753 bp	DNA
LOCUS			linear
DEFINITION			PAT 27-AUG-2002
ACCESSION	BD014992		
VERSION	BD014992.1		GI:22555799
KEYWORDS	JP 2001169788-A/2.		
SOURCE	unidentified		
ORGANISM	unclassified.		
REFERENCE	1 (bases 1 to 753)		
AUTHORS	KennethKunehito, N., Sahm, H., Eggering, L. and Pfefferle, W.		
TITLE	Nucleotide sequence encoding the discharge of branched amino acid, method of isolating the same and utilization thereof		
JOURNAL	Patent: JP 2001169788-A 2-26-JUN-2001;		
COMMENT	DEBUSA HUELS AG, FORSCHUNGSZENTRUM JUELICH GMBH		
	OS Corynebacterium glutamicum ATCC14752		
	PN JP 2001169788-A/2		
	PD 26-JUN-2001		
	PF 24-OCT-2000 JP 2000324315		
	PR 27-OCT-1999 DE 19551708.8		
	PI NICOLE KENNERUKUMEHITO, HERMANN SAHM, IOTHAER EGGERING, WALTER PFEFFERLE		
	PC C12N15/09, C07K14/34, C12N1/21, C12P13/06, C12P13/06, C12P13/08//		
	PC (C12N15/09, C12R1.15), (C12N1/21, C12R1.15), (C12P13/06, C12R1.15),		
	PC C12N15/00, C12R1.15)		
	PC C12N15/00, C12R1.15)		
	CC brnF		
	CC key		
	CC key		
	FT CDS		
FEATURES			
source	1..753	Location/Qualifiers	
	location="unidentified"		
	/organism="genomic DNA"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:32644"		
ORIGIN			
Query Match	100.0%; Score 753; DB 6; Length 753;		
Best Local Similarity	100.0%; Pred. No. 9.ee-185;		
Matches	753; Conservative 0; Mismatches 0; Indels 0; Gaps 0		
Qy	1	GTGCAAAAACGCAAGATTCATTCAGGCTCGAGGCTGTGGCCATCCAGGCAAGCCCTG	60
Db	1	GTGCAAAAACGCAAGATTCATTCAGGCTCGAGGCTGTGGCCATCCAGGCAAGCCCTG	60
Qy	61	GAAACAGATGATTAAGGTTATCGGCGCTACGAAATCGCGCAAGGCTCTAAAACCTCCCT	120
Db	61	GAAACAGATGATTAAGGTTATCGGCGCTACGAAATCGCGCAAGGCTCTAAAACCTCCCT	120
Qy	121	GCTGAGGTTGGGCGCATGATCCCGATGGTATGGTATGGTCTCTTGTTATTCGAATAC	180
Db	121	GCTGAGGTTGGGCGCATGATCCCGATGGTATGGTATGGTCTCTTGTTATTCGAATAC	180

QY 181 GGCTACGATGATGGGAGCCCACTGTTTCCGGCTGATTTTCGGGGCTCCACCGAA 240
 DB 181 GGCTACGATGATGGGAGCCCACTGTTTCCGGCTGATTTTCGGGGCTCCACCGAA 240
 QY 241 ATGCTGATGATGCGCTGTTGTTGGGCGGAGGCGCCCTGGGCGCATGCGGCTTACCA 300
 DB 241 ATGCTGATGATGCGCTGTTGTTGGGCGGAGGCGCCCTGGGCGCATGCGGCTTACCA 300
 QY 301 TTGCTGATGATGCGCTGTTGTTGGGCGGAGGCGCCCTGGGCGCATGCGGCTTACCA 360
 DB 301 TTGCTGATGATGCGCTGTTGTTGGGCGGAGGCGCCCTGGGCGCATGCGGCTTACCA 360
 QY 361 AACCCATTGCGCGCTTCTATTGCGGTTTTCGGCTTATGACGAAAGCTTACGCTCACT 420
 DB 361 AACCCATTGCGCGCTTCTATTGCGGTTTTCGGCTTATGACGAAAGCTTACGCTCACT 420
 QY 421 GGGGCGAGGCGCGGAGGCTGTCGGGCGGAGCTTATGCAATGCAATGAGCTTTCAC 480
 DB 421 GGGGCGAGGCGCGGAGGCTGTCGGGCGGAGCTTATGCAATGCAATGAGCTTTCAC 480
 QY 481 TCTTACTGAGTATTCGGCGGCTCTCAACCGAGTGGCGATGCGAGAGTTGATTCCTTTGAA 540
 DB 481 TCTTACTGAGTATTCGGCGGCTCTCAACCGAGTGGCGATGCGAGAGTTGATTCCTTTGAA 540
 QY 541 ATTAAAGGCTGAGATTCGCGCTTTCCTCTCTTTTTCAGAGCTTTCAGATTCCTTTCG 600
 DB 541 ATTAAAGGCTGAGATTCGCGCTTTCCTCTCTTTTTCAGAGCTTTCAGATTCCTTTCG 600
 QY 601 CGAAGCAAAAAGAGATTCCTTCTCTGCTGCTGCGAGTTTGAAGCTTCAACATTCCTCT 660
 DB 601 CGAAGCAAAAAGAGATTCCTTCTCTGCTGCTGCGAGTTTGAAGCTTCAACATTCCTCT 660
 QY 661 GTGCTATTCAGATGAGGCTTTCCTTTCGCGGCTGCTGATCTTCTTGGGTCTGTGACC 720
 DB 661 GTGCTATTCAGATGAGGCTTTCCTTTCGCGGCTGCTGATCTTCTTGGGTCTGTGACC 720
 QY 721 ATCCGATCTTCTTCTTGGGAAAGCTGCTTAA 753
 DB 721 ATCCGATCTTCTTCTTGGGAAAGCTGCTTAA 753
 RESULT 5
 BD162490 753 bp DNA linear PAT 17-JAN-2003
 LOCUS Novel polynucleotide.
 DEFINITION BD162490
 ACCESSION BD162490.1 GI:27686248
 VERSION JP 2002191370-A/289.
 KEYWORDS unidentified
 SOURCE unidentified
 ORGANISM unclassified.
 REFERENCE 1 (bases 1 to 753)
 AUTHORS Nakagawa, S., Mizoguchi, H., Ando, S., Hayashi, M., Ochiai, K.,
 TITLE Yokoi, H., Tateishi, N., Senoo, A., Ikeda, M. and Ozaki, A.
 JOURNAL Novel polynucleotide
 Patent: JP 2002191370-A 289 09-JUL-2002;
 KYOMA HAKKO KOGYO CO LTD
 OS Corynebacterium glutamicum
 PN JP 2002191370-A/289
 PD 09-JUL-2002
 PF 15-DEC-2000 JP 2000405096
 PI SATOSHI NAKAGAWA, HIROSHI MIZOGUCHI, SEIKO ANDO, MIKIO HAYASHI,
 PI KEIKO OCHIAI,
 PI HARUHIRO YOKOI, NAKO TATEISHI, AKIHIRO SENOO, MASATO IKEDA, AKIO
 PI OZAKI
 PC C12N15/09, C12N15/09, C07K14/34, C07K16/12, C07K16/40, C12M1/00, PC
 C12N1/15,
 PC C12N1/19, C12N1/21, C12N5/10, C12N9/00, C12N9/02, C12P7/40, C12P13/
 PC 04, C12P13/08,
 PC C12P19/00, C12P19/34, C12P21/02, C12Q1/37, C12Q1/68, G01N33/53, PC
 G01N33/56,
 PC G01N33/569, G01N33/68, G01N37/00, C12P21/08, C12N1/21, C12R1/15,
 PC C12N1/21, C12R1/13, C12N1/21, C12R1/01, C12P13/08, C12R1/15,

PC C12N15/00,
 PC C12N5/00, C12N15/00
 CC Novel polynucleotide
 FH Key Location/Qualifiers
 FT source 1..753
 FT /organism='Corynebacterium glutamicum'.
 FEATURES
 source
 1..753
 /organism="unidentified"
 /mol_type="genomic DNA"
 /db_xref="taxon:3264"
 ORIGIN
 Query Match 100.0%; Score 753; DB 6; Length 753;
 Best Local Similarity 100.0%; Pred No. 9, 8e-185; Indels 0; Gaps 0;
 Matches 753; Conservative 0; Mismatches 0;
 QY 1 GTCAAAAAGAGAGATTCATTCAGCTGAGAGTGTCCCATCAAGGAGCCCTG 60
 DB 1 GTCAAAAAGAGAGATTCATTCAGCTGAGAGTGTCCCATCAAGGAGCCCTG 60
 QY 61 GAACGAGATATTAAGTTATCGGCGCTACGAAATCGCGAAGTCTAATAAAGCTCCCT 120
 DB 61 GAACGAGATATTAAGTTATCGGCGCTACGAAATCGCGAAGTCTAATAAAGCTCCCT 120
 QY 121 GCTGAGGTTTGGGCGATGACCCGATTTGATTTGCTGTTGCTCTTGTATTCAATAC 180
 DB 121 GCTGAGGTTTGGGCGATGACCCGATTTGATTTGCTGTTGCTCTTGTATTCAATAC 180
 QY 181 GGCTACGATGATGGGAGCCCACTGTTTCCGGCTGATTTTCGGGCTCCACCGAA 240
 DB 181 GGCTACGATGATGGGAGCCCACTGTTTCCGGCTGATTTTCGGGCTCCACCGAA 240
 QY 241 ATGCTGATGATGCGCTGTTGTTGGGCGGAGGCGCCCTGGGCGCATGCGGCTTACCA 300
 DB 241 ATGCTGATGATGCGCTGTTGTTGGGCGGAGGCGCCCTGGGCGCATGCGGCTTACCA 300
 QY 301 TTGCTGATGATGCGCTGTTGTTGGGCGGAGGCGCCCTGGGCGCATGCGGCTTACCA 360
 DB 301 TTGCTGATGATGCGCTGTTGTTGGGCGGAGGCGCCCTGGGCGCATGCGGCTTACCA 360
 QY 361 AACCCATTGCGCGCTTCTATTGCGGTTTTCGGCTTATGACGAAAGCTTACGCTCACT 420
 DB 361 AACCCATTGCGCGCTTCTATTGCGGTTTTCGGCTTATGACGAAAGCTTACGCTCACT 420
 QY 421 GGGGCGAGGCGCGGAGGCTGTCGGGCGGAGCTTATGCAATGCAATGAGCTTTCAC 480
 DB 421 GGGGCGAGGCGCGGAGGCTGTCGGGCGGAGCTTATGCAATGCAATGAGCTTTCAC 480
 QY 481 TCTTACTGAGTATTCGGCGGCTCTCAACCGAGTGGCGATGCGAGAGTTGATTCCTTTGAA 540
 DB 481 TCTTACTGAGTATTCGGCGGCTCTCAACCGAGTGGCGATGCGAGAGTTGATTCCTTTGAA 540
 QY 541 ATTAAAGGCTGAGATTCGCGCTTTCCTCTCTTTTTCAGAGCTTTCAGATTCCTTTCG 600
 DB 541 ATTAAAGGCTGAGATTCGCGCTTTCCTCTCTTTTTCAGAGCTTTCAGATTCCTTTCG 600
 QY 601 CGAAGCAAAAAGAGATTCCTTCTCTGCTGCTGCGAGTTTGAAGCTTCAACATTCCTCT 660
 DB 601 CGAAGCAAAAAGAGATTCCTTCTCTGCTGCTGCGAGTTTGAAGCTTCAACATTCCTCT 660
 QY 661 GTGCTATTCAGATGAGGCTTTCCTTTCGCGGCTGCTGATCTTCTTGGGTCTGTGACC 720
 DB 661 GTGCTATTCAGATGAGGCTTTCCTTTCGCGGCTGCTGATCTTCTTGGGTCTGTGACC 720
 QY 721 ATCCGATCTTCTTCTTGGGAAAGCTGCTTAA 753
 DB 721 ATCCGATCTTCTTCTTGGGAAAGCTGCTTAA 753
 RESULT 5
 AR391953 1271 bp DNA linear PAT 18-DEC-2003
 LOCUS AR391953

DEFINITION Sequence 1 from patent US 6613545.
 AR391953
 VERSION AR391953.1 GI:40115724
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE
 1 (bases 1 to 1271)
 Unclassified.
 AUTHORS Kernerhnecht,N., Sahm,H., Eggeling,L. and Pfeifferle,W.
 TITLE Nucleotide sequences coding for the export of branched amino acids, process for the isolation thereof and use thereof
 JOURNAL Patent: US 6613545-A 1 02-SEP-2003;
 FEATURES
 source Location/Qualifiers
 1. 1271
 /organism="unknown"
 /mol_type="genomic DNA"

Query Match 100.0%; Score 753; DB 6; Length 1271;
 Best Local Similarity 100.0%; Pred. No. 1e-184;
 Matches 753; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTGCAAAAACGCAAGATTGATTCAGAGCTGAGAGTGTGCCATCCAGGACGCTG 60
 DB 101 GTGCAAAAACGCAAGATTGATTCAGAGCTGAGAGTGTGCCATCCAGGACGCTG 160
 QY 61 GAACCAAGATGATAAAGTTATCGGCGCTACGAATCGCGAAGGTCTAAAACTCCCTT 120
 DB 161 GAACCAAGATGATAAAGTTATCGGCGCTACGAATCGCGAAGGTCTAAAACTCCCTT 220
 QY 121 GCTGCAAGTTTGGGCAATGATCCGATGGTATTCGCTTGTCTCTTGATTATCAATAC 180
 DB 221 GCTGCAAGTTTGGGCAATGATCCGATGGTATTCGCTTGTCTCTTGATTATCAATAC 280
 QY 181 GGCTACGAATGATGGGCAAGCCCACTGTTTCCGGCTGATTTTGGCGGGCTCCACCGAA 240
 DB 281 GGCTACGAATGATGGGCAAGCCCACTGTTTCCGGCTGATTTTGGCGGGCTCCACCGAA 340
 QY 241 ATGCTGATGATGCGCTGCTGTTGTGGGCGCAGCGCCCTGAGCGCATGCGCTCACCA 300
 DB 341 ATGCTGATGATGCGCTGCTGTTGTGGGCGCAGCGCCCTGAGCGCATGCGCTCACCA 400
 QY 301 TTGCTGATGATGCGCTGCTGTTGTGGGCGCAGCGCTGATTTTCAATCCGCTGATGCTCAA 360
 DB 401 TTGCTGATGATGCGCTGCTGTTGTGGGCGCAGCGCTGATTTTCAATCCGCTGATGCTCAA 460
 QY 361 AACCCGATGCGCGCTTCTATTCGATTCGATTCGCGCTTATGACGAAGACCTACAGTCACT 420
 DB 461 AACCCGATGCGCGCTTCTATTCGATTCGATTCGCGCTTATGACGAAGACCTACAGTCACT 520
 QY 421 GGGGCGAAGCGCGCAGAGCTGTGCGGCGTGGCACTTATCTCAATGCAATTAACGTTTAC 480
 DB 521 GGGGCGAAGCGCGCAGAGCTGTGCGGCGTGGCACTTATCTCAATGCAATTAACGTTTAC 580
 QY 481 TCTCTACTGGGTAATTCGCGCGTCTCAACGAGTGGCGATGCGAGATTGATTCCTTTTGA 540
 DB 581 TCTCTACTGGGTAATTCGCGCGTCTCAACGAGTGGCGATGCGAGATTGATTCCTTTTGA 640
 QY 541 ATTAAGGCGCTCGAGTTGCGCGCTTGTCTCTCTTTGTGACGCTGACTTGGATTCTGCG 600
 DB 641 ATTAAGGCGCTCGAGTTGCGCGCTTGTCTCTCTTTGTGACGCTGACTTGGATTCTGCG 700
 QY 601 CGAAGCAAAAAGCAGATCCCTCTCTGCTGCTGCGAGTTTGAAGTTCACCATTCCTCTT 660
 DB 701 CGAAGCAAAAAGCAGATCCCTCTCTGCTGCTGCGAGTTTGAAGTTCACCATTCCTCTT 760
 QY 661 GTGTAATTCAGATGAGGCGCTTATTTGGCGGCTGCTGATCTTCTTGGGCTGCTTTGAC 720
 DB 761 GTGTAATTCAGATGAGGCGCTTATTTGGCGGCTGCTGATCTTCTTGGGCTGCTTTGAC 820
 QY 721 ATCCGATCTCTCTTGGGAAAGGCTGCTAAA 753
 DB 821 ATCCGATCTCTCTTGGGAAAGGCTGCTAAA 853

RESULT 7
 AR391956
 LOCUS 1271 bp DNA linear PAT 18-DEC-2003
 DEFINITION
 Sequence 6 from patent US 6613545.
 AR391956
 VERSION AR391956
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE
 1 (bases 1 to 1271)
 Unclassified.
 AUTHORS Kernerhnecht,N., Sahm,H., Eggeling,L. and Pfeifferle,W.
 TITLE Nucleotide sequences coding for the export of branched chain amino acids, process for the isolation thereof and use thereof
 JOURNAL Patent: US 6613545-A 6 02-SEP-2003;
 FEATURES
 source Location/Qualifiers
 1. 1271
 /organism="unknown"
 /mol_type="genomic DNA"

Query Match 100.0%; Score 753; DB 6; Length 1271;
 Best Local Similarity 100.0%; Pred. No. 1e-184;
 Matches 753; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTGCAAAAACGCAAGATTGATTCAGAGCTGAGAGTGTGCCATCCAGGACGCTG 60
 DB 101 GTGCAAAAACGCAAGATTGATTCAGAGCTGAGAGTGTGCCATCCAGGACGCTG 160
 QY 61 GAACCAAGATGATAAAGTTATCGGCGCTACGAATCGCGAAGGTCTAAAACTCCCTT 120
 DB 161 GAACCAAGATGATAAAGTTATCGGCGCTACGAATCGCGAAGGTCTAAAACTCCCTT 220
 QY 121 GCTGCAAGTTTGGGCAATGATCCGATGGTATTCGCTTGTCTCTTGATTATCAATAC 180
 DB 221 GCTGCAAGTTTGGGCAATGATCCGATGGTATTCGCTTGTCTCTTGATTATCAATAC 280
 QY 181 GGCTACGAATGATGGGCAAGCCCACTGTTTCCGGCTGATTTTGGCGGGCTCCACCGAA 240
 DB 281 GGCTACGAATGATGGGCAAGCCCACTGTTTCCGGCTGATTTTGGCGGGCTCCACCGAA 340
 QY 241 ATGCTGATGATGCGCTGCTGTTGTGGGCGCAGCGCCCTGAGCGCATGCGCTCACCA 300
 DB 341 ATGCTGATGATGCGCTGCTGTTGTGGGCGCAGCGCCCTGAGCGCATGCGCTCACCA 400
 QY 301 TTGCTGATGATGCGCTGCTGTTGTGGGCGCAGCGCTGATTTTCAATCCGCTGATGCTCAA 360
 DB 401 TTGCTGATGATGCGCTGCTGTTGTGGGCGCAGCGCTGATTTTCAATCCGCTGATGCTCAA 460
 QY 421 GGGGCGAAGCGCGCAGAGCTGTGCGGCGTGGCACTTATCTCAATGCAATTAACGTTTAC 480
 DB 521 GGGGCGAAGCGCGCAGAGCTGTGCGGCGTGGCACTTATCTCAATGCAATTAACGTTTAC 580
 QY 481 TCTCTACTGGGTAATTCGCGCGTCTCAACGAGTGGCGATGCGAGATTGATTCCTTTTGA 540
 DB 581 TCTCTACTGGGTAATTCGCGCGTCTCAACGAGTGGCGATGCGAGATTGATTCCTTTTGA 640
 QY 541 ATTAAGGCGCTCGAGTTGCGCGCTTGTCTCTCTTTGTGACGCTGACTTGGATTCTGCG 600
 DB 641 ATTAAGGCGCTCGAGTTGCGCGCTTGTCTCTCTTTGTGACGCTGACTTGGATTCTGCG 700
 QY 601 CGAAGCAAAAAGCAGATCCCTCTCTGCTGCTGCGAGTTTGAAGTTCACCATTCCTCTT 660
 DB 701 CGAAGCAAAAAGCAGATCCCTCTCTGCTGCTGCGAGTTTGAAGTTCACCATTCCTCTT 760
 QY 661 GTGTAATTCAGATGAGGCGCTTATTTGGCGGCTGCTGATCTTCTTGGGCTGCTTTGAC 720

Db 761 GTGTAATTCAGAGCCCTATTGGGCGCTGATCTTCTTGAGTAC 820

Qy 721 ATCCGTAATCTTCTTGGGAAAGCTGTAA 753

Db 821 ATCCGTAATCTTCTTGGGAAAGCTGTAA 853

RESULT 8

AX137709 1271 bp DNA linear PAT 30-MAY-2001

LOCUS Sequence 1 from Patent EP1096010.

DEFINITION AX137709

ACCESSION AX137709

VERSION AX137709.1 GI:14273886

KEYWORDS

SOURCE

ORGANISM

Corynebacterium glutamicum

Corynebacterium glutamicum

Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;

Corynebacteriaceae; Corynebacterium.

REFERENCE

1 Kemerhocht N., Eggeling L., Sahm H. and Pfeffertle W.

Nucleotide sequences coding for branched-chain amino acids export

protein, method for isolating them and their use

Patent: EP 1096010-A 1 02-MAY-2001;

Degussa AG (DE) ; FORSCHUNGSZENTRUM JUELICH GMBH (DE)

LOCATION/Qualifiers

1..1271

/organism="Corynebacterium glutamicum"

/mol_type="unassigned DNA"

/db_xref="taxon:1718"

/note="ATCC14752"

101..856

/note="unnamed protein product; brnF"

/codon_start=1

/transl_table=1

/protein_id="CAC39923.1"

/db_xref="GI:14273887"

/db_xref="EMBL:CAC39923"

/db_xref="REFSEQ:NM_001080481.1"

/translation="MQKQEIHSLSLEVSPPSKALBPDDKGYRREYIAQGLKTSIAAGL

GVNPIGIAFGLLVQYGYEMWAPLPSGLIFAGSTEMLVIALVGAAPLGAIALTL

VNFRHVFVAFSPPLHVNPIARFVSFALIDBAVYATARPAQMSAMRLISMQIAFH

SVWFGGLGVAIAELIPEIKGLPEALCSFVTLTLDSCRTKQIPSLIAGLSFTI

ALVPIPGALPAALLIFGLITIRYPLGKAK"

853..1179

/note="unnamed protein product; brnE"

/codon_start=1

/transl_table=1

/protein_id="CAC39924.1"

/db_xref="GI:14273888"

/db_xref="EMBL:CAC39924"

/db_xref="REFSEQ:NM_001080482.1"

/translation="MTDPSGTLIVAVCAVTPARVAPLPIKRIEPSQVGMAM

WMPAGLIALITASTFRSNAIDKTLTFGLIAVITVAAILGGRITLSVAGTIVFV

GLVNL"

ORIGIN

Query Match 100.0%; Score 753; DB 6; Length 1271;

Best Local Similarity 100.0%; Pred. No. 1e-164;

Matches 753; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GTGCAAAAAACGCAAGATTCATTCAAGCTGAGGTGTGCCATCCAAAGCAGCCCTG 60

Db 101 GTGCAAAAAACGCAAGATTCATTCAAGCTGAGGTGTGCCATCCAAAGCAGCCCTG 160

Qy 61 GAACCAATATTAAGGTTATCGCGCTAGCAATCGCGAAGTCTAAAAACCTCCCTT 120

Db 161 GAACCAATATTAAGGTTATCGCGCTAGCAATCGCGAAGTCTAAAAACCTCCCTT 220

Qy 121 GCTGCAAGTTGGCATGTACCCGATGATGATGCGTTGCTCTTTGTTATCAATAC 180

Db 221 GCTGCAAGTTGGCATGTACCCGATGATGATGCGTTGCTCTTTGTTATCAATAC 280

Qy 181 GGCTACGAATGGTGGAGGCCCACTGTTTTCGGGCTGATTTTCGGGCTCCACCGAA 240

Db 281 GGCTACGAATGGTGGAGGCCCACTGTTTTCGGGCTGATTTTCGGGCTCCACCGAA 340

Qy 241 ATGCTGATCAATGCGCCCTGTTGAGGAGGAGGCGCCCTGAGGCGCATGGGCGTCAACACA 300

Db 341 ATGCTGATCAATGCGCCCTGTTGAGGAGGAGGCGCCCTGAGGCGCATGGGCGTCAACACA 400

Qy 301 TTGCTGATCAATGCGCCCTGTTGAGGAGGAGGCGCCCTGAGGCGCATGGGCGTCAACACA 360

Db 401 TTGCTGATCAATGCGCCCTGTTGAGGAGGAGGCGCCCTGAGGCGCATGGGCGTCAACACA 460

Qy 361 AACCCATTCGCGCTTCTATTTGCGTTTTCGGCTTATGAGCAAGCCATGAGTCACT 420

Db 461 AACCCATTCGCGCTTCTATTTGCGTTTTCGGCTTATGAGCAAGCCATGAGTCACT 520

Qy 421 GCGGCAAGCCGCGAGCTGTGCGGCGTGCAGTATCTCATGCAATAGCGTTTAC 480

Db 521 GCGGCAAGCCGCGAGCTGTGCGGCGTGCAGTATCTCATGCAATAGCGTTTAC 580

Qy 481 TCTACTGAGTATTCGCGCTTCTACCGAGTGCAGTATGCAATGATTTCTTTGAA 540

Db 581 TCTACTGAGTATTCGCGCTTCTACCGAGTGCAGTATGCAATGATTTCTTTGAA 640

Qy 541 ATTAAGGCGCTCGAGTTCGCGCTTCTCTCTTGTGCAAGCTGATTTGATTCCTGC 600

Db 641 ATTAAGGCGCTCGAGTTCGCGCTTCTCTCTTGTGCAAGCTGATTTGATTCCTGC 700

Qy 601 CGAACAAAAACGATCCCTTCTCTGCTGCTGCGAGTTTGAAGCTTACCAATTCCTT 660

Db 701 CGAACAAAAACGATCCCTTCTCTGCTGCTGCGAGTTTGAAGCTTACCAATTCCTT 760

Qy 661 GTGTAATTCAGGTCAGGCGCTTATTTGGGCGGTGTGATTTCTTGGTGTGAGC 720

Db 761 GTGTAATTCAGGTCAGGCGCTTATTTGGGCGGTGTGATTTCTTGGTGTGAGC 820

Qy 721 ATCCGTAATCTTCTTGGGAAAGCTGTAA 753

Db 821 ATCCGTAATCTTCTTGGGAAAGCTGTAA 853

RESULT 9

AX137714 1271 bp DNA linear PAT 30-MAY-2001

LOCUS Sequence 6 from Patent EP1096010.

DEFINITION AX137714

ACCESSION AX137714

VERSION AX137714.1 GI:14273893

KEYWORDS

SOURCE

ORGANISM

Corynebacterium glutamicum

Corynebacterium glutamicum

Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;

Corynebacteriaceae; Corynebacterium.

REFERENCE

1 Kemerhocht N., Eggeling L., Sahm H. and Pfeffertle W.

Nucleotide sequences coding for branched-chain amino acids export

protein, method for isolating them and their use

Patent: EP 1096010-A 6 02-MAY-2001;

Degussa AG (DE) ; FORSCHUNGSZENTRUM JUELICH GMBH (DE)

LOCATION/Qualifiers

1..1271

/organism="Corynebacterium glutamicum"

/mol_type="unassigned DNA"

/db_xref="taxon:1718"

/note="ATCC13032"

101..856

/note="unnamed protein product; brnF"

/codon_start=1

/transl_table=1

/protein_id="CAC39927.1"

/db_xref="GI:14273894"

/db_xref="EMBL:CAC39927"

/db_xref="REFSEQ:NM_001080483.1"

/translation="MQKQEIHSLSLEVSPPSKALBPDDKGYRREYIAQGLKTSIAAGL

GVNPIGIAFGLLVQYGYEMWAPLPSGLIFAGSTEMLVIALVGAAPLGAIALTL

VNFRHVFVAFSPPLHVNPIARFVSFALIDBAVYATARPAQMSAMRLISMQIAFH

SVWFGGLGVAIAELIPEIKGLPEALCSFVTLTLDSCRTKQIPSLIAGLSFTI

ALVPIPGALPAALLIFGLITIRYPLGKAK"

FEATURES

source

CDS
853..1179
/note="unnamed protein product; brnE"
/codon_start=1
/transl_table=11
/protein_id="CAC39928.1"
/db_xref="GI:14273895"
/db_xref="REFSEQ:CAC39928"
/translation="MTDFSCILVAVCAVITPALRAVPELLIKPRESOVKMM
MMPAGILAILASTFRSNAIDIKLITFGLIAVAITVAHLGRRRLISVAGTIVEV
GLVNLV"

ORIGIN

Query Match 100.0%; Score 753; DB 6; Length 1271;
Best Local Similarity 100.0%; Pred. No. 1e-184;
Matches 753; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 GTGCAAAAACCGAAGATTCATTCAGAGCTGAGGTGCGCATCCAGAGAGCCCTG 60
101 GTGCAAAAACCGAAGATTCATTCAGAGCTGAGGTGCGCATCCAGAGAGCCCTG 160
61 GAACCAATGATTAAGTTATCGCGCTACGAATCGCGCAAGGTCTAAAACCTCCCTT 120
161 GAACCAATGATTAAGTTATCGCGCTACGAATCGCGCAAGGTCTAAAACCTCCCTT 220
121 GGTGAGGTTTGGGAGATGTAACCGATTTGATTTGCTCTGTTATTCATATC 180
221 GGTGAGGTTTGGGAGATGTAACCGATTTGATTTGCTCTGTTATTCATATC 280
181 GGCTACGAATGATGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240
281 GGCTACGAATGATGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 340
241 AGGCTGATATGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
341 AGGCTGATATGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 400
301 TTGCTGATGATTCGAGAGATTCATTCAGAGCTGAGGTGCGCATCCAGAGAGCC 360
401 TTGCTGATGATTCGAGAGATTCATTCAGAGCTGAGGTGCGCATCCAGAGAGCC 460
361 AACCCATGAGGAG 420
461 AACCCATGAGGAG 520
421 GGCGCAG 480
521 GGCGCAG 580
481 TCTCTAGGATTAAGTTATCGCGCTACGAATCGCGCAAGGTCTAAAACCTCCCTT 540
581 TCTCTAGGATTAAGTTATCGCGCTACGAATCGCGCAAGGTCTAAAACCTCCCTT 640
541 ATTAAGGAGCTGAGAGTTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
641 ATTAAGGAGCTGAGAGTTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 700
601 CGAAGCAAAAAGAGATTCCTCTCTGCTGCTGCGAGGTTAGCTTCAACATTCCTCT 660
701 CGAAGCAAAAAGAGATTCCTCTCTGCTGCTGCGAGGTTAGCTTCAACATTCCTCT 760
661 GTGCTAATTCAGAGTTCAGAGCTGATTTGAGGAGAGAGAGAGAGAGAGAGAGAG 720
761 GTGCTAATTCAGAGTTCAGAGCTGATTTGAGGAGAGAGAGAGAGAGAGAGAGAG 820
721 ATCCGATCTTCTTCTGAGAAAGGCTGCTAA 753
821 ATCCGATCTTCTTCTGAGAAAGGCTGCTAA 853

RESULT 10
BD014991 1271 bp DNA linear PAT 27-AUG-2002
LOCUS BD014991
DEFINITION Nucleotide sequence encoding the discharge of branched amino acid,

method of isolating the same and utilization thereof.
ACCESSION BD014991
VERSION BD014991.1 GI:2255798
KEYWORDS JP 2001169788-A/1.
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1 (bases 1 to 1271)
AUTHORS Kennerkumehito, N., Sahm, H., Eggering, L. and Pfeifferle, W.
TITLE Nucleotide sequence encoding the discharge of branched amino acid,
METHOD OF isolating the same and utilization thereof
PATENT: JP 2001169788-A 1 26-JUN-2001
JOURNAL DEGUSA HUELS AG, FORSCHUNGSZENTRUM JUELICH GMBH
COMMENT OS Corynebacterium glutamicum ATCC14752
FN JP 2001169788-A/1
PD 26-JUN-2001
PF 24-OCT-2000 JP 2000324315
PR 27-OCT-1999 DE 19951708.8
PI NICOLE KENNERKUMEHITO, HERMANN SAHM, LOTHAR EGGERING, WALTER PI
PFEIFFERLE
PC C12N15/09, C07K14/34, C12N1/21, C12P13/06, C12P13/06, C12P13/08//
PC (C12N15/09, C12R1:15), (C12N1/21, C12R1:15), (C12P13/06, C12R1:15),
PC (C12N15/00, C12R1:15)
CC brnE
CC brnE
FH Key Location/Qualifiers
FT gene (101)..(853)
FT Location/Qualifiers (853)..(1176).
1..1271
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"

ORIGIN
Query Match 100.0%; Score 753; DB 6; Length 1271;
Best Local Similarity 100.0%; Pred. No. 1e-184;
Matches 753; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 GTGCAAAAACCGAAGATTCATTCAGAGCTGAGGTGCGCATCCAGAGAGCCCTG 60
101 GTGCAAAAACCGAAGATTCATTCAGAGCTGAGGTGCGCATCCAGAGAGCCCTG 160
61 GAACCAATGATTAAGTTATCGCGCTACGAATCGCGCAAGGTCTAAAACCTCCCTT 120
161 GAACCAATGATTAAGTTATCGCGCTACGAATCGCGCAAGGTCTAAAACCTCCCTT 220
121 GGTGAGGTTTGGGAGATGTAACCGATTTGATTTGCTCTGTTATTCATATC 180
221 GGTGAGGTTTGGGAGATGTAACCGATTTGATTTGCTCTGTTATTCATATC 280
181 GGCTACGAATGATGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240
281 GGCTACGAATGATGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 340
241 AGGCTGATATGAGGAG 300
341 AGGCTGATATGAGGAG 400
301 TTGCTGATGATTCGAGAGATTCATTCAGAGCTGAGGTGCGCATCCAGAGAGCC 360
401 TTGCTGATGATTCGAGAGATTCATTCAGAGCTGAGGTGCGCATCCAGAGAGCC 460
361 AACCCATGAGGAG 420
461 AACCCATGAGGAG 520
421 GGCGCAG 480
521 GGCGCAG 580
481 TCTCTAGGATTAAGTTATCGCGCTACGAATCGCGCAAGGTCTAAAACCTCCCTT 540

Db 581 TCCACTGGGATTCGCGGCTCTACCGAGTGGCGATGCAAGTGAATTCCTTTGAA 640

Qy 541 ATTAAGGCGCTGAGATTGCGCCCTTTGCTCTCTTTGTACGCTGACTTTGGAATTCCTGC 600

Db 641 ATTAAGGCGCTGAGATTGCGCCCTTTGCTCTCTTTGTACGCTGACTTTGGAATTCCTGC 700

Qy 601 CGAAGAAAAGACATCCCTCTGCTGCTGCGAGGTTTGAAGTTCAACCATTCCTT 660

Db 701 CGAAGAAAAGACATCCCTCTGCTGCTGCGAGGTTTGAAGTTCAACCATTCCTT 760

Qy 661 GTGTAATTCAGATCAGGCGCTTATTTGCGGCGCTGATCTTCTTGGGTCTGTGAC 720

Db 761 GTGTAATTCAGATCAGGCGCTTATTTGCGGCGCTGATCTTCTTGGGTCTGTGAC 820

Qy 721 ATCCGTAATCTTCTTCTTGGGAAAGCTGCTAA 753

Db 821 ATCCGTAATCTTCTTCTTGGGAAAGCTGCTAA 853

RESULT 11

LOCUS BD014994 1271 bp DNA linear PAT 27-AUG-2002

DEFINITION Nucleotide sequence encoding the discharge of branched amino acid, method of isolating the same and utilization thereof.

ACCESSION BD014994.1 GI:22555801

VERSION JP 2001169788-A/4.

KEYWORDS unclassified

SOURCE unclassified

ORGANISM unclassified

REFERENCE 1 (bases 1 to 1271)
Kemerunknecht,N., Sahm,H., Eggering,L. and Pfeifferle,W.
Nucleotide sequence encoding the discharge of branched amino acid, method of isolating the same and utilization thereof
Patent: JP 2001169788-A 4 26-JUN-2001.

JOURNAL DEBUSA HUBLS AG,FORSCHUNGSZENTRUM JUELICH GMBH

COMMENT OS Corynebacterium glutamicum ATCC13032
PN JP 2001169788-A/4
PD 26-JUN-2001
PF 24-OCT-2000 JP 2000324315
PR 27-OCT-1999 DE 19951708.8
PI NICOLE KEMERUNKNECHT,HERMANN SAHM,LOTHAR EGGERING,WALTER PI PFEIFFERLE

PC C12N15/09,C07K14/34,C12N1/21,C12P13/06,C12P13/06,C12P13/06//
PC (C12N15/09,C12R1:15), (C12N1/21,C12R1:15), (C12P13/06,C12R1:15),
PC (C12N15/00,
CC C12N15/00,C12R1:15)
CC brnF
CC key
FH Key
FT gene
FT Location/Qualifiers
1..1271
/organism="unclassified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"

ORIGIN

Query Match 100.0%; Score 753; DB 6; Length 1271;
Best Local Similarity 100.0%; Pred. No. 1e-184;
Matches 753; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GTGAAAAAAGCGAAGATTCATTCAGGCTGAGGTGCGCATCCAAAGAGCCCTG 60

Db 101 GTGAAAAAAGCGAAGATTCATTCAGGCTGAGGTGCGCATCCAAAGAGCCCTG 160

Qy 61 GAACCAATGATTAAGGTTATCGCGCTACGAAATCCGCGAAGGTCTTAAACCTCCCT 120

Db 161 GAACCAATGATTAAGGTTATCGCGCTACGAAATCCGCGAAGGTCTTAAACCTCCCT 220

Qy 121 GCTGAGGTTTGGGCAATGACCCGATTTGATTCGTTGGTCTCTTGTATTCAATAC 180

Db 221 GCTGAGGTTTGGGCAATGACCCGATTTGATTCGTTGGTCTCTTGTATTCAATAC 280

Qy 181 GGTACGAATGATGGGACACCCCACTGTTTTCGCGCTGATTTTGGCGGCTCCACCGAA 240

Db 281 GGTACGAATGATGGGACACCCCACTGTTTTCGCGCTGATTTTGGCGGCTCCACCGAA 340

Qy 241 ATGCTGATATGCGCCCTCTGTTGAGCGAGCGCCCTTGGGCGCATTCGCTGACACA 300

Db 341 ATGCTGATATGCGCCCTCTGTTGAGCGAGCGCCCTTGGGCGCATTCGCTGACACA 400

Qy 301 TTGCTGTAATCTTCCGCAAGTATCTATGGGTTTCAATCCCGTGAATGGTCAAA 360

Db 401 TTGCTGTAATCTTCCGCAAGTATCTATGGGTTTCAATCCCGTGAATGGTCAAA 460

Qy 361 AACCCATTGCGCGTTTCTATTCGTTTTCGCGTTATCGAGAAAGCTTACGACATCACT 420

Db 461 AACCCATTGCGCGTTTCTATTCGTTTTCGCGTTATCGAGAAAGCTTACGACATCACT 520

Qy 421 GCGGCGAGCGCGAGCTGCTGCGGCTGCGGCACTTATCTCAATGCAATAGGTTTCA 480

Db 521 GCGGCGAGCGCGAGCTGCTGCGGCTGCGGCACTTATCTCAATGCAATAGGTTTCA 580

Qy 481 TCCACTGGGATTCGCGGCTCTACCGAGTGGCGATGCAAGTGAATTCCTTTGAA 540

Db 581 TCCACTGGGATTCGCGGCTCTACCGAGTGGCGATGCAAGTGAATTCCTTTGAA 640

Qy 541 ATTAAGGCGCTGAGATTGCGCCCTTTGCTCTCTTTGTACAGCTGACTTTGATTCCTGC 600

Db 641 ATTAAGGCGCTGAGATTGCGCCCTTTGCTCTCTTTGTACAGCTGACTTTGATTCCTGC 700

Qy 601 CGAAGAAAAGACATCCCTCTGCTGCTGCGAGTGTGAGCTTCAACCATTCCTT 660

Db 701 CGAAGAAAAGACATCCCTCTGCTGCTGCGAGTGTGAGCTTCAACCATTCCTT 760

Qy 661 GTGTAATTCAGATCAGGCGCTTATTTGCGGCGCTGATCTTCTTGGGTCTGTGAC 720

Db 761 GTGTAATTCAGATCAGGCGCTTATTTGCGGCGCTGATCTTCTTGGGTCTGTGAC 820

Qy 721 ATCCGTAATCTTCTTCTTGGGAAAGCTGCTAA 753

Db 821 ATCCGTAATCTTCTTCTTGGGAAAGCTGCTAA 853

RESULT 12

LOCUS AP454053 2105 bp DNA linear BCT 27-JUN-2002

DEFINITION Corynebacterium glutamicum brnE (brnE), brnF (brnF), and Lrp-1-like regulator (lrp) genes, complete cds.

ACCESSION AP454053

VERSION AP454053.1 GI:21311379

KEYWORDS Corynebacterium glutamicum

SOURCE Corynebacterium glutamicum

ORGANISM Corynebacterium glutamicum

REFERENCE 1 (bases 1 to 2105)
Kemerunknecht,N., Sahm,H., Yen,M.R., Patek,M., Sailer Jr,M.H. Jr. and Eggering,L.
Export of L-isoleucine from Corynebacterium glutamicum: a two-gene-encoded member of a new translocator family
J. Bacteriol. 184 (14), 3947-3956 (2002)

JOURNAL MEDLINE 22077265

PUBMED 12081967

REFERENCE 2 (bases 1 to 2105)
Kemerunknecht,N., Eggering,L. and Sahm,H.
Direct Submission
Submitted (07-NOV-2001) IBI-1, Forschungszentrum, Leo-Brandt Str., Juelich 52425, Germany

JOURNAL Location/Qualifiers
1..2105
/organism="Corynebacterium glutamicum"
/mol_type="genomic DNA"

gene	/db_xref="taxon:1718"	
CDS	complement(385..711)	
	/gene="brnE"	
	complement(385..711)	
	/gene="brnE"	
	/note="LIV-E transporter component 1; branched-chain amino acid export protein"	
	/codon_start=1	
	/transl_table=11	
	/product="BrnE"	
	/protein_id="AA04685.1"	
	/db_xref="GI:21311380"	
	/translation="MTDPSCLLVAVCAVITPAAVPELILKPLRESQFVKMAWMPACILALITASTRSNAIDKITPGLIVAVITVAHLGRRITLISVAGTIVEGLVNLJF"	
gene	complement(708..1463)	
CDS	/gene="brnF"	
	complement(708..1463)	
	/gene="brnF"	
	/note="LIV-E transporter component 2; branched-chain amino acid export protein"	
	/codon_start=1	
	/transl_table=11	
	/product="BrnF"	
	/protein_id="AA04686.1"	
	/db_xref="GI:21311381"	
	/translation="MOKTOEHSLEVSPSKALPPDKGYRVEIAGLTKSLAGLGRPIGIAFGLVIOGYBWWAPLPSGLIFRGSFEMVILVAVGASARLISQIARHVNFRVFAFSPLEHVNPIARFVSFALIDEAVIAPAPAGASARLISQIARHSYVWFGALGVAIAELIPFELKGFALCSFVTLTLDSCRTKQIPSLLAGSFTLALVIRPGALFEALILFLILIRYFLGRKAK"	
gene	1538..1993	
CDS	/gene="ltp"	
	1538..1993	
	/gene="ltp"	
	/note="Ltp"	
	/codon_start=1	
	/transl_table=11	
	/product="Ltp-like regulator"	
	/protein_id="AA04687.1"	
	/db_xref="GI:21311382"	
	/translation="MKLSDIRATIAESANRISNLALADKHTLTPPCRRORLEAEGLIGSADHHPVWNRGEFVTVDSITVDFSNFSSVAQHDVEVLEHLRFGSPDYFVRIGVADLEAYEQFLSHIQVPIAKISRRFAMKVKPARPV"	
ORIGIN		
Query Match	100.0%; Score 753; DB 1; Length 2105;	
Best Local Similarity	100.0%; Pred. No. 1e-184;	
Matches /53; Conservative	0; Mismatches 0; Indels 0; Gaps 0;	
QY	1 GTGCAAAAAGCAGAGATTCATTCAGCCTGAGAGTGTGCGCATCCAGGAGCCCTG 60	
DB	1463 GTGCAAAAAGCAGAGATTCATTCAGCCTGAGAGTGTGCGCATCCAGGAGCCCTG 1404	
QY	61 GAACCGAGATTAAGCTATGCGGCTGCGAATCGGACGAGCTGTAACCTCCCT 120	
DB	1403 GAACCGAGATTAAGCTATGCGGCTGCGAATCGGACGAGCTGTAACCTCCCT 1344	
QY	121 GGTGAGGTTTGGGCGATGACCGATGGATGGCTTTGCTCTCTGGTTATTCATAC 180	
DB	1343 GGTGAGGTTTGGGCGATGACCGATGGATGGCTTTGCTCTCTGGTTATTCATAC 1284	
QY	181 GGTAGGAATGGTGGCAGCCCACTGTTTCCGCTCATTTTGGCGGCTCCACCGAA 240	
DB	1283 GGTAGGAATGGTGGCAGCCCACTGTTTCCGCTCATTTTGGCGGCTCCACCGAA 1224	
QY	241 ATGCTGATCGCCTCGTTGTGGGCGAGCCGCTGAGCGCATCGCGTCAACCA 300	
DB	1223 ATGCTGATCGCCTCGTTGTGGGCGAGCCGCTGAGCGCATCGCGTCAACCA 1164	
QY	301 TTGCTGGTGAACCTCGGCGAGATTCATGCTTTTCATTCGCGTGCATGTGTCAAA 360	
DB	1163 TTGCTGGTGAACCTCGGCGAGATTCATGCTTTTCATTCGCGTGCATGTGTCAAA 1104	
QY	361 AACCCATGCGCGTTTCATTCGCTTTTGGGCTTATGAGAGACCTTACGAGTCACT 420	
DB	1103 AACCCATGCGCGTTTCATTCGCTTTTGGGCTTATGAGAGACCTTACGAGTCACT 1044	
QY	421 GCGGCGAGCGCCGAGAGCTGTGCGGCGAGCACTTATCTCAATGCAATAGGTTTCA 480	
DB	1043 GCGGCGAGCGCCGAGAGCTGTGCGGCGAGCACTTATCTCAATGCAATAGGTTTCA 984	
QY	481 TCTTACTGGATTCGCGCGCTCTACCGAGATGCGCATGCGAGAGTTTCTTTGAA 540	
DB	983 TCTTACTGGATTCGCGCGCTCTACCGAGATGCGCATGCGAGAGTTTCTTTGAA 924	
QY	541 ATTAAGGCGCTGAGTTGCGCCTTGTCTCTCTTGTGACGCTGACTTGAATCTGCG 600	
DB	923 ATTAAGGCGCTGAGTTGCGCCTTGTCTCTCTTGTGACGCTGACTTGAATCTGCG 864	
QY	601 CGAAGCAAAAAGAGATCCCTTCTCTGCTGCTGCGAGTTTGAGCTTCAACATGCTCT 660	
DB	863 CGAAGCAAAAAGAGATCCCTTCTCTGCTGCTGCGAGTTTGAGCTTCAACATGCTCT 804	
QY	661 GTGTAATTCAGAGTCAAGGCTTATTTGGCGGCTGCTATCTTTGGGTCTGTGACC 720	
DB	803 GTGTAATTCAGAGTCAAGGCTTATTTGGCGGCTGCTATCTTTGGGTCTGTGACC 744	
QY	721 ATCCGTAATTCCTTTCTTGGAGAAAGCTGCTAAA 753	
DB	743 ATCCGTAATTCCTTTCTTGGAGAAAGCTGCTAAA 711	
RESULT 13		
AP005274	340000 bp DNA linear BCT 08-AUG-2002	
LOCUS	Corynebacterium glutamicum ATCC 13032 DNA, complete genome, section 1/10	
DEFINITION	AP005274 Corynebacterium glutamicum ATCC 13032 DNA, complete genome, section 1/10	
ACCESSION	AP005274 BA000036	
KEYWORDS	AP005274.1 GI:21322764	
SOURCE	Corynebacterium glutamicum ATCC 13032	
ORGANISM	Corynebacterium glutamicum ATCC 13032	
REFERENCE	Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales; Corynebacteriaceae; Corynebacteriineae; Corynebacterium.	
AUTHORS	1 Nakagawa, S.	
TITLE	Complete genomic sequence of Corynebacterium glutamicum ATCC 13032	
JOURNAL	Unpublished	
REFERENCE	2 (bases 1 to 340000)	
AUTHORS	Nakagawa, S.	
TITLE	Direct Submission	
JOURNAL	Submitted (04-MAY-2002) Satoshi Nakagawa, Kyowa Hakko Kogyo Co. Ltd., Tokyo Research Laboratories; 3-6-6, Asahi-machi, Machida, Tokyo 194-8533, Japan (E-mail: snakagawa@xanaden.com, Tel:81-44-829-3831, Fax:81-44-813-1651)	
COMMENT	This sequence is deposited by collaboration of Kyowa Hakko Kogyo Co. Ltd. And Kitasato University.	
FEATURES	Location/Qualifiers	
source	1..340000	
	/organism="Corynebacterium glutamicum ATCC 13032"	
	/mol_type="genomic DNA"	
	/strain="ATCC 13032"	
	/db_xref="taxon:196627"	
	/note="ATCC 13032"	
gene	1..1575	
	/gene="Cgl10001"	
CDS	1..1575	
	/gene="Cgl10001"	
	/note="PF00308: Bacterial dnaA protein	
	TIGR00362: dnaA: chromosomal replication initiator protein	
	dnaA"	
	/codon_start=1	
	/transl_table=11	
	/product="Atpase involved in DNA replication initiation"	
	/protein_id="BAB7394.1"	

```
/db_xref="GI:21322765"
/translation="MSQNSSSLLETRWQVADLTLLSQADSGPDLTPPQAVINLT
KRIALVDGIATVSTPNAAKNTIENDLGLATRVLSLRKGRFSFLAVSEPEQIIPET
PAQOEFKYQDPAPVISNKAPOQVEGGEASTDQWERTHSAPAPHPAPLADPE
PELATPQRIPRETPANPNREVSINPKYFFSEFVIGPNRFANAAVAASPAKFN
PLFISGSGELKTHLHVAAGYAOELQPLRIKYVSEETNDYINSRDROEFTFR
RYRNDILAVDDIOFLAGKEGTQEEFFFNALHOADKOIILSDRPPKOLITLEDRL
TRFRGGILTDIOPDLETRIILBKQOTDGTWDRVLELIASRFSISRILEGAL
IKVASSILINQPIDKEMALVALRDILPEPEMETATAPIMEVTAYEFESIVDTLGA
GKTRAVAHARQAMTLCRELTDMSLPKIGDVGSGKHITWTADRKIKQEMTERKDTY
DEIQQLTQIKSRGRN"
complement(1594..1920)
/gene="Cg10002"
/complement(1594..1920)
/gene="Cg10002"
/codon_start=1
/translation="Hypothetical protein"
/protein_id="BAB97395.1"
/db_xref="GI:21322766"
/translation="MIVDNFVYLMIDFKYRILSPYCADVSPAYPOFLGSCAGGKFRG
ITPPVTKLELQVCNTAVNKKVDNVSAFISGYPQNCCTPKLQKQPSDEGSSRAA
EINFTV"
2292..3476
/gene="Cg10003"
2292..3476
/gene="Cg10003"
/EC_number="2.7.7.7"
/note="pf00712:DNA polymerase III beta subunit, N-terminal
domain
pf02767:DNA polymerase III beta subunit, central domain
pf02768:DNA polymerase III beta subunit, C-terminal domain
TIGR00663:dnan: DNA polymerase III, beta subunit"
/codon_start=1
/translation="
/product="DNA polymerase III beta subunit"
/protein_id="BAB97396.1"
/db_xref="GI:21322767"
/translation="MESQVSEFVAREDIVTAVAYANRLPTKPTQPIRAMLITADD
EGLEAGYDVYSTRVLSAQSQGRIVAGKLSLSEITGSLPNKVFVRIDIGSAFV
TCSSRFELPLIDYPMPLPKLPAGTGSINAKLFTSAVSQVSAAGXDSLPMLTGV
SMEIVGQINIAATDFRLALRTFEWANDPELNVLLIPARTLLDNASLDSGLNDS
IDIVAGTGQVGRSGEGVYTDNRERTTLADPNNIAPLPKEHTAIASVEIAPLV
DAIRVSIYAEPAACIVLHRSQCVILTATGACAGAEETLPCATGKSLTIAAPGY
LKDGLSVPTSRVAFGFTSPSPAINIPEPEKPSANENGITQTPDTTYTILMVRLL
PG"
3585..4769
/gene="Cg10004"
3585..4769
/gene="Cg10004"
/note="TIGR00611:recf: Recf protein"
/codon_start=1
/translation="
/product="Recombinational DNA repair ApPase"
/protein_id="BAB97397.1"
/db_xref="GI:21322768"
/translation="MHIRSLDELDRYSWPELKVDLEPGITVFIRGNFGKTNIVEAIG
YLAHLSHRVSSDAPLIVRAHENAARSAVANVGRRLAHLIKPRAAGASLNPYK
RTPRELIGVKTTLFAPEDIALYKGPBARERRLIDITLITROPMAQVADYDKLTK
RNMLKTAITALKRGGTGEBGAALSTLTDMDQALRLGAEVAAAPALINLGPRTY
EAYTTIAPESRPAVNTKTTIDQSLQSFSEFDGIIIEATLLELAKRPRELERGSSL
VGPRHDVDTLMDOPAKGVSAGETWSFALSRLIAEFNLKSDGTDPTLLIDVSESE
LDAGREKLVIGIAQVEVQVLTAAVDDIPENLKVLTAQHTVTVQDTGTGRISILDV
Op"
4766..5302
/gene="Cg10005"
4766..5302
/gene="Cg10005"
/codon_start=1
/translation="Hypothetical protein"
/protein_id="BAB97398.1"
/db_xref="GI:21322769"
/translation="MTDPIEQAFERIRAEAMRRNGSVPLDNKNDAFRRPAPRGVEK
RKGRASGADGROKRYRGAGEISGLVANKELIRRGKQDIAGWTSNMEELVGAKIA
QHTRVEMIDKRLPITCDSTAMATNLARMQRLQVIAKVGPNITTELRITGPQAS
WRKGPVAVKRGGRDITYG"
5354..7489
/gene="Cg10006"
5354..7489
/gene="Cg10006"
/EC_number="5.99.1.3"
/note="pf00204:DNA topoisomerase II (N-terminal region)
pf00986:DNA gyrase B subunit, carboxyl terminus
pf01751:toprim domain
pf02518:Histidine kinase", DNA gyrase B-, phytochrome-like
ATPase
TIGR01055:pare Gneg: DNA topoisomerase IV, B subunit
TIGR01058:pare Gpos: DNA topoisomerase IV, B subunit
TIGR01059:gyrB: DNA gyrase, B subunit"
/codon_start=1
/translation="
/product="DNA gyrase (topoisomerase II) B subunit"
/protein_id="BAB97399.1"
/db_xref="GI:21322770"
/translation="MGGTWTGPKRVKMKRLYRISSECSLKVANTENHNDASSITILE
GLEVRKRPQMYTIGTGPRLHILWEVDNVSVDAMAGATVEVTLLEDGVOYVD
DGRGIPYDMHPSGAPTYQVYMTOLHAGKXDSIYAVSGHGVGISVYNALSTRVEA
DIRLHGHWYQNFESVYDVLISGNARQGTITIRWPDALITETIEPDETSRLQ
EMAFLNKGLITITLDNNAITDELELEALQGSTALELDEIDNETLVEETDAK
KPKREKKKIIFYPNGLIEDVYHANSKTNIHISYIFSAKGDHVEVAAQWNSYX
ESVTFPANTINTEREGTHEEGRSALSTLNNRVARBKLIKEXANLTGDDCEGISA
VIRVAGDPOFEGQTKTLGNTREKSTVQGMANEHIGHMLEANPAKTYINKAVASA
QARLAAKARDIVRKSAITDGLPGCLAKRDEKSESLYIVEDSGSAGKSGSD
SMRQALIPILGKILNFKALDLYLKNAEQALITLGCIIHBPINKLRHYKIYIM
ADADVQGHATLTLLEFRFMDLVABGVITLAQPLVCLKRQREPGAYISDERD
EQLNEGIAARCKINKDQIGRYKGLGEMNASLEMTTMDPTVILRRVDITDQRADE
LFSILMGDDVVAARSFTRNAKQVFLDI"
7830..8798
/gene="Cg10007"
7830..8798
/gene="Cg10007"
/codon_start=1
/translation="Hypothetical protein"
/protein_id="BAB97400.1"
/db_xref="GI:21322771"
/translation="MIQTYLIKIOSLKWAQREMLKTLINDTLGHVTSFEPTTAHP
LPDGSSTPVOIKMASDNDSQVYMLMPFGMGYVYRPLAALAKAGHVAIGLRQGG
QSSAKARSKSQMGYTHLASVDFPQIAPAKKALDLEBGHMRFLSHMSGQICLPA
RPEAKKYNLRAIRFGVAGSFPRTFSPKWKARLGLAVLLGGIGGHI VGGPQVLAK
DLVQYGRQSGTHMRKWRFRKHSLLDITLQODINVEVMKVASIPIITFSCPDDECP
QASIDGLASFVPAQIQMIEIPALGHNWARBESYVQLFLBOAL"
complement(8795..9466)
/gene="Cg10008"
complement(8795..9466)
/gene="Cg10008"
/codon_start=1
/translation="
/product="Hypothetical membrane protein"
/protein_id="BAB97401.1"
/db_xref="GI:21322772"
/translation="MSNPNKRPDLDAQRFSLPRMAMFVGFIFMSPAIGITVHAGIG
TTTISLPYVNTASGSLSKWTTIYRNGNMIQIYLRQFQKQMLVOLIMLFLPFR
LCDLSLQGTIAQDNTFVAMITVIVSTILMSISVFFCYVNPNTIFIGSGEIVSALYVK
FPNVEFGIMQIYDMTVFSVAALISWTMGDLGVRGTVPAFFIGFVRQMKLYL
RSIGH"
complement(9471..9914)
/gene="Cg10009"
complement(9471..9914)
/gene="Cg10009"
/codon_start=1
/translation="Hypothetical protein"
/protein_id="BAB97402.1"
/db_xref="GI:21322773"
```

```

/translation="MTSLRLDLVSTNDAIADMSAFIEETVKGSGISGIALKGMG
AATKIDSDIVAKSRRLPEIADSLDGLMOEXONGTADPAHLFANSSRALAIIIS
VADKNAESINVPGLKVKYGVGKAARKYIEGLPRLIGLIENMQ"
complement(10104..11177)
/gene="Cg10010"
/complement(10104..11177)
/gene="Cg10010"
/codon_start=1
/transl_table=11
/product="Hypothetical protein"
/protein_id="BAH97403.1"
/db_xref="GI:21322774"
/translation="MEMLAQRPLPHITAPHLVAVRPHQIRATDADRTSLDRLELS

```

Query Match 100.0%; Score 753; DB 1; Length 340000;
 Best Local Similarity 100.0%; Pred. No. 1.2e-184;
 Matches 753; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 GTGCAAAAAACGCAAGATTCATTCAAGCTGAGGTGTCCCATCCAGGAGCCCTG 60
DB 276829 GTGCAAAAAACGCAAGATTCATTCAAGCTGAGGTGTCCCATCCAGGAGCCCTG 276888
QY 61 GAACCAAGATTAAGTTATCGGCGCTCGAAGATGCGCAAGGTCTAAAACTCCCTT 120
DB 276889 GAACCAAGATTAAGTTATCGGCGCTCGAAGATGCGCAAGGTCTAAAACTCCCTT 276948
QY 121 GCTGAGGTTGGGCAATGACCGATGATGCGTTGATGCTCTTGTTATCAATAC 180
DB 276949 GCTGAGGTTGGGCAATGACCGATGATGCGTTGATGCTCTTGTTATCAATAC 277008
QY 181 GGTAGAAATGTTGGGCAAGCCCACTGTTTCCGGCTGATTTTCCGGGGCTCCACCGA 240
DB 277009 GGTAGAAATGTTGGGCAAGCCCACTGTTTCCGGCTGATTTTCCGGGGCTCCACCGA 277068
QY 241 ATGCTGGTCAATCGCCCTCGTTGTGGGCGAGCGCCCTGGGCGCATCCGCTCACCA 300
DB 277069 ATGCTGGTCAATCGCCCTCGTTGTGGGCGAGCGCCCTGGGCGCATCCGCTCACCA 277128
QY 301 TTGCTGTGAATTCGCGCAAGTATTCATGCGTTTCAATCCCGCTGATGTGCTGAA 360
DB 277129 TTGCTGTGAATTCGCGCAAGTATTCATGCGTTTCAATCCCGCTGATGTGCTGAA 277188
QY 361 AACCCATTGCGCGTTTCAATTCGTTTTCGCGCTTATCGACGAAGCTACGAGTCACT 420
DB 277189 AACCCATTGCGCGTTTCAATTCGTTTTCGCGCTTATCGACGAAGCTACGAGTCACT 277248
QY 421 GCGGCGAAGCGCGAGGCTGTGGCGGTGGCGCATTAATCTCAATGCAAAATAGCGTTTAC 480
DB 277249 GCGGCGAAGCGCGAGGCTGTGGCGGTGGCGCATTAATCTCAATGCAAAATAGCGTTTAC 277308
QY 481 TCCTAATGAGTATTCGCGCGGTCTCACCGAGTGGCGCATGCGAAGTTGATTCCTTTGAA 540
DB 277309 TCCTAATGAGTATTCGCGCGGTCTCACCGAGTGGCGCATGCGAAGTTGATTCCTTTGAA 277368
QY 541 ATTAAAGGCTCGAGTTGCGCCCTTGTCTCTCTTTGTCAAGCTGACTTTGATTCCTGC 600
DB 277369 ATTAAAGGCTCGAGTTGCGCCCTTGTCTCTCTTTGTCAAGCTGACTTTGATTCCTGC 277428
QY 601 CGAAGCAAAAAGCAAGATCCCTTCTGTCGTCGCGAGTTGAGTTCACCAATGCTCT 660
DB 277429 CGAAGCAAAAAGCAAGATCCCTTCTGTCGTCGCGAGTTGAGTTCACCAATGCTCT 277488
QY 661 GTGGTAATTCAGAGTCAAGCCCTATTGTCGCGCTGCTGTAATCTCTTGAGTGTGACC 720
DB 277489 GTGGTAATTCAGAGTCAAGCCCTATTGTCGCGCTGCTGTAATCTCTTGAGTGTGACC 277548
QY 721 ATCCGTAATTCCTTTCTTGGAAGGCTGCTAAA 753
DB 277549 ATCCGTAATTCCTTTCTTGGAAGGCTGCTAAA 277581

```

RESULT 14
 AX120085

```

LOCUS AX120085 349980 bp DNA linear PAT 11-MAY-2001
DEFINITION Sequence 1 from Patent EP1108790.
ACCESSION AX120085 AX114121
VERSION AX120085.1 GI:114036800
KEYWORDS
ORGANISM
SOURCE
Corynebacterium glutamicum
Corynebacterium glutamicum
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacterineae; Corynebacteriaceae; Corynebacterium.
REFERENCE
1 Nakagawa, S., Mizoguchi, H., Ando, S., Hayashi, M., Ochiai, K.,
Yokoi, H., Tateishi, N., Senoh, A., Ikeda, M. and Ozaki, A.
Novel polynucleotides
Patent: EP 1108790-A 1 20-JUN-2001;
JOURNAL KYOMA HAKKO KOGYO CO., LTD. (JP)
FEATURES
source
1..349980
/organism="Corynebacterium glutamicum"
/mol_type="genomic DNA"
/db_xref="taxon:1718"
/note="Seq 1 to long (3,309,400) split in 11, seq 1
0.000,001 0.349,980 349,980"

```

ORIGIN

Query Match 100.0%; Score 753; DB 6; Length 349980;
 Best Local Similarity 100.0%; Pred. No. 1.2e-184;
 Matches 753; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 GTGCAAAAAACGCAAGATTCATTCAAGCTGAGGTGTCCCATCCAGGAGCCCTG 60
DB 276829 GTGCAAAAAACGCAAGATTCATTCAAGCTGAGGTGTCCCATCCAGGAGCCCTG 276888
QY 61 GAACCAAGATTAAGTTATCGGCGCTCGAAGATGCGCAAGGTCTAAAACTCCCTT 120
DB 276889 GAACCAAGATTAAGTTATCGGCGCTCGAAGATGCGCAAGGTCTAAAACTCCCTT 276948
QY 121 GCTGAGGTTGGGCAATGACCGATGATGCGTTGATGCTCTTGTTATCAATAC 180
DB 276949 GCTGAGGTTGGGCAATGACCGATGATGCGTTGATGCTCTTGTTATCAATAC 277008
QY 181 GGTAGAAATGTTGGGCAAGCCCACTGTTTCCGGCTGATTTTCCGGGGCTCCACCGA 240
DB 277009 GGTAGAAATGTTGGGCAAGCCCACTGTTTCCGGCTGATTTTCCGGGGCTCCACCGA 277068
QY 241 ATGCTGGTCAATCGCCCTCGTTGTGGGCGAGCGCCCTGGGCGCATCCGCTCACCA 300
DB 277069 ATGCTGGTCAATCGCCCTCGTTGTGGGCGAGCGCCCTGGGCGCATCCGCTCACCA 277128
QY 301 TTGCTGTGAATTCGCGCAAGTATTCATGCGTTTCAATCCCGCTGATGTGCTGAA 360
DB 277129 TTGCTGTGAATTCGCGCAAGTATTCATGCGTTTCAATCCCGCTGATGTGCTGAA 277188
QY 361 AACCCATTGCGCGTTTCAATTCGTTTTCGCGCTTATCGACGAAGCTACGAGTCACT 420
DB 277189 AACCCATTGCGCGTTTCAATTCGTTTTCGCGCTTATCGACGAAGCTACGAGTCACT 277248
QY 421 GCGGCGAAGCGCGAGGCTGTGGCGGTGGCGCATTAATCTCAATGCAAAATAGCGTTTAC 480
DB 277249 GCGGCGAAGCGCGAGGCTGTGGCGGTGGCGCATTAATCTCAATGCAAAATAGCGTTTAC 277308
QY 481 TCCTAATGAGTATTCGCGCGGTCTCACCGAGTGGCGCATGCGAAGTTGATTCCTTTGAA 540
DB 277309 TCCTAATGAGTATTCGCGCGGTCTCACCGAGTGGCGCATGCGAAGTTGATTCCTTTGAA 277368
QY 541 ATTAAAGGCTCGAGTTGCGCCCTTGTCTCTCTTTGTCAAGCTGACTTTGATTCCTGC 600
DB 277369 ATTAAAGGCTCGAGTTGCGCCCTTGTCTCTCTTTGTCAAGCTGACTTTGATTCCTGC 277428
QY 601 CGAAGCAAAAAGCAAGATCCCTTCTGTCGTCGCGAGTTGAGTTCACCAATGCTCT 660
DB 277429 CGAAGCAAAAAGCAAGATCCCTTCTGTCGTCGCGAGTTGAGTTCACCAATGCTCT 277488
QY 661 GTGGTAATTCAGAGTCAAGCCCTATTGTCGCGCTGCTGTAATCTCTTGAGTGTGACC 720

```


CC This invention describes a novel isolated polynucleotide (1) containing
CC at least one sequence that (i) is 70% identical with a sequence that
CC encodes at least one of 251 amino acid (aa) (3) or 108 aa (5)
CC polypeptides; (ii) encodes a polypeptide at least 70% identical with (3)
CC or (5); (iii) is the complement of (i) or (ii); or (iv) contains at least
CC 15 consecutive bases from (i)-(iii); The invention also describes (a)
CC protein derived from sequences of 753 bp (2) or 324 bp (4); (b)
CC Corynebacterium microorganisms, especially Corynebacterium, transformed with
CC one or more (1), where these are replicative DNA; (c) production of

CC branched-chain L-aa by fermentation of coryneform bacteria in which the
CC bms and/or bms genes (or equivalent sequences) are amplified,
CC especially overexpressed; and (d) method for isolating the bms and/or
CC bms genes. (1) is used for transformation of coryneform bacteria being
CC used for fermentative production of branched-chain amino acids,
CC specifically leucine, isoleucine and valine, which are useful in medicine
CC and animal nutrition. (1) can also be used as source of primers and
CC probes for isolation of related sequences. Transformation with (1)
CC increases yield of branched-chain amino acids. This sequence encodes the
CC Corynebacterium glutamicum ATCC 14752 bms protein described in the
CC method of the invention

XX Sequence 753 BP; 140 A; 211 C; 194 G; 208 T; 0 U; 0 Other;

Query Match 100.0%; Score 753; DB 4; Length 753;
Best Local Similarity 100.0%; Pred. No. 2e-223;
Matches 753; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTGCAAAAAACGCAAGATTCATTCAAGCTGAGAGTGTCCGATCCAGAGGAGCCCTG 60
DB 1 GTGCAAAAAACGCAAGATTCATTCAAGCTGAGAGTGTCCGATCCAGAGGAGCCCTG 60
QY 61 GAACAGATGATTAAGATTATCGAGCTACGAATCGCGCAAGGTCTTAAACCTCCCTT 120
DB 61 GAACAGATGATTAAGATTATCGAGCTACGAATCGCGCAAGGTCTTAAACCTCCCTT 120
QY 121 GCTGCAAGTTTGGGCATGATACCCGATGATGATGCTGTTGCTGTTGTTATTCATAC 180
DB 121 GCTGCAAGTTTGGGCATGATACCCGATGATGATGCTGTTGCTGTTGTTATTCATAC 180
QY 181 GGTGCAAGATGATGAGGAGGAGCCGACATGTTTCCGAGCTGATTTCCGAGGCTCCAGCA 240
DB 181 GGTGCAAGATGATGAGGAGGAGCCGACATGTTTCCGAGCTGATTTCCGAGGCTCCAGCA 240
QY 241 ATGCTGATCGCCCTCGTTGTTGGGCGAGCGCCCTGAGGCGCATCGGCTCACACA 300
DB 241 ATGCTGATCGCCCTCGTTGTTGGGCGAGCGCCCTGAGGCGCATCGGCTCACACA 300
QY 301 TTGCTGATGAACTCCGCGACGATTCATGCGTTTCAATCCGCTGATGATGATGATGAT 360
DB 301 TTGCTGATGAACTCCGCGACGATTCATGCGTTTCAATCCGCTGATGATGATGATGAT 360
QY 361 AACCCATGCGCCCTTCTATTGCGTTTTCGCTTATCGACGAGCGCTACGAGTCACT 420
DB 361 AACCCATGCGCCCTTCTATTGCGTTTTCGCTTATCGACGAGCGCTACGAGTCACT 420
QY 421 GCGGCGAGGCGCGAGGCTGTGCGGCTGCGGCTGATTCATGCAATGCAATGCAATGCA 480
DB 421 GCGGCGAGGCGCGAGGCTGTGCGGCTGCGGCTGATTCATGCAATGCAATGCAATGCA 480
QY 481 TCCTACTGATGATTCGCGGCTGATCGAGGAGTGCATGCGAGTGTGATTCCTTTGAA 540
DB 481 TCCTACTGATGATTCGCGGCTGATCGAGGAGTGCATGCGAGTGTGATTCCTTTGAA 540
QY 541 ATTAAAGGCTTCGAGTTCGCTTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 600
DB 541 ATTAAAGGCTTCGAGTTCGCTTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 600
QY 601 CGAAGCAAAAAAGCATCCCTTCTCTGCTGCTCGAGGTTGAGTTGACATTCATTCCTT 660
DB 601 CGAAGCAAAAAAGCATCCCTTCTCTGCTGCTCGAGGTTGAGTTGACATTCATTCCTT 660
QY 661 GTGATTAATTCAGGTCAGGCTTATTTGGGCGGCTGCTGATTCCTTGGGCTGTGAC 720
DB 661 GTGATTAATTCAGGTCAGGCTTATTTGGGCGGCTGCTGATTCCTTGGGCTGTGAC 720
QY 721 ATCCGATCTCTCTCTGAGGAAAGGCTGCTAAA 753
DB 721 ATCCGATCTCTCTCTGAGGAAAGGCTGCTAAA 753

RESULT 2
AAH65254

ID AAH65254 standard; DNA; 753 BP.
AC AAH65254;
XX
XX
XX 26-SEP-2001 (first entry)
DE C glutamicum coding sequence fragment SEQ ID NO: 289.
XX
XX Corynebacterium; amino acid synthesis; vitamin; saccharide;
KM organic acid synthesis; ds.
XX
XX
OS Corynebacterium glutamicum.
XX
XX EPI108790-A2.
XX
XX 20-JUN-2001.
XX
XX 18-DEC-2000; 2000EP-00127688.
XX
XX 16-DEC-1999; 99JP-00377484.
XX
XX 07-APR-2000; 2000JP-00159162.
XX
XX 03-AUG-2000; 2000JP-00280988.
XX
XX (KYOWA) KYOWA HAKKO KOGYO KK.
XX
XX Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;
XX Tateishi N, Senoh A, Ikeda M, Ozaki A;
XX WPI; 2001-376931/40.
XX P-PSDB; AAG90035.
XX
XX Novel polynucleotides derived from Corynebacterium bacteria, for identifying
XX mutation point of a gene, measuring expression of a gene, analyzing
XX expression profile or pattern of a gene and identifying homologous gene.
XX
XX Claim 8; SEQ ID NO 289; 246pp + Sequence listing; English.
XX
XX The present invention provides a number of nucleotide and protein
XX sequences from the Corynebacterium glutamicum. These
XX are useful for identifying the mutation point of a gene derived from a
XX mutant of corynebacterium, measuring expression amount and analysing
XX the expression profile or expression pattern of a gene derived from
XX Corynebacterium, and identifying a homologue of a gene derived from
XX Corynebacterium. Corynebacterium bacteria are useful for producing amino
XX acids, nucleic acids, vitamins, saccharides and organic acids,
XX particularly L-lysine. The present sequence is a nucleic acid described
XX in the exemplification of the invention. Note: The sequence data for this
XX patent did not form part of the printed specification but was obtained
XX in electronic format directly from the European Patent Office

SEQ Sequence 753 BP; 140 A; 211 C; 194 G; 208 T; 0 U; 0 Other;
Query Match 100.0%; Score 753; DB 5; Length 753;
Best Local Similarity 100.0%; Pred. No. 2e-223;
Matches 753; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTGCAAAAAACGCAAGATTCATTCAAGCTGAGAGTGTCCGATCCAGAGGAGCCCTG 60
DB 1 GTGCAAAAAACGCAAGATTCATTCAAGCTGAGAGTGTCCGATCCAGAGGAGCCCTG 60
QY 61 GAACAGATGATTAAGATTATCGAGCTACGAATCGCGCAAGGTCTTAAACCTCCCTT 120
DB 61 GAACAGATGATTAAGATTATCGAGCTACGAATCGCGCAAGGTCTTAAACCTCCCTT 120
QY 121 GCTGCAAGTTTGGGCATGATACCCGATGATGATGCTGTTGCTGTTGTTATTCATAC 180
DB 121 GCTGCAAGTTTGGGCATGATACCCGATGATGATGCTGTTGCTGTTGTTATTCATAC 180
QY 181 GGTGCAAGATGATGAGGAGGAGCCGACATGTTTCCGAGCTGATTTCCGAGGCTCCAGCA 240
DB 181 GGTGCAAGATGATGAGGAGGAGCCGACATGTTTCCGAGCTGATTTCCGAGGCTCCAGCA 240
QY 241 ATGCTGATCGCCCTCGTTGTTGGGCGAGCGCCCTGAGGCGCATCGGCTCACACA 300

Db 241 ATGCTGATCATGCGCTGCTGTGTGGGCGCAGCCGCCGCGCATGCGCTCACACACA 300
 Qy 301 TTGCTGGGAACTTCCGCGCAGATTTCTATGCGTTTTCATTCGCGCGCATGCTGCAAA 360
 Db 301 TTGCTGGGAACTTCCGCGCAGATTTCTATGCGTTTTCATTCGCGCGCATGCTGCAAA 360
 Qy 361 AACCCCATTTGCCCGTTTCTATTCGCTTTCGCGCTTATGCAAGAGCTTACGAGTCACT 420
 Db 361 AACCCCATTTGCCCGTTTCTATTCGCTTTCGCGCTTATGCAAGAGCTTACGAGTCACT 420
 Qy 421 GCGGCGAGGCGCGCAGGCTGTGCGCGCGCGAGCTTATCTCAATGCAATGCGTTTCAAC 480
 Db 421 GCGGCGAGGCGCGCAGGCTGTGCGCGCGCGAGCTTATCTCAATGCAATGCGTTTCAAC 480
 Qy 481 TCCTACTGGGATTTGCGCGGCTCTCACCGAGTGGCGATCGAGAGTTGATTCCTTTGAA 540
 Db 481 TCCTACTGGGATTTGCGCGGCTCTCACCGAGTGGCGATCGAGAGTTGATTCCTTTGAA 540
 Qy 541 ATTAAGGCGCTGAGTTCGCGCTTTCGCTCTCTTTTGTACGCTGACTTGGATTCCTGC 600
 Db 541 ATTAAGGCGCTGAGTTCGCGCTTTCGCTCTCTTTTGTACGCTGACTTGGATTCCTGC 600
 Qy 601 CGAAGCAAAAGACAGATCCCTTCTGCTGCTGCGAGGTTGAGGTTGACGATTCCTT 660
 Db 601 CGAAGCAAAAGACAGATCCCTTCTGCTGCTGCGAGGTTGAGGTTGACGATTCCTT 660
 Qy 661 GTGGTAATTCAGGTCAGGCGCTTATTTGCGGCGCTGATCTTCTTGGATCTGTGAC 720
 Db 661 GTGGTAATTCAGGTCAGGCGCTTATTTGCGGCGCTGATCTTCTTGGATCTGTGAC 720
 Qy 721 ATCCGCTACTTCTTCTTCTTGGGAAAAGGCTCTTAA 753
 Db 721 ATCCGCTACTTCTTCTTGGGAAAAGGCTCTTAA 753

RESULT 3

AAH21109 standard; DNA; 1271 BP.

AAH21109;

05-SEP-2001 (first entry)

C. glutamicum DNA encoding brnF and brnE.

L-amino acid production; brnF, brnE; branched-chain amino acid;

Corynebacterium; leucine; isoleucine; valine; medicine;

animal nutrition; ds.

Corynebacterium glutamicum.

Key Location/Qualifiers

FT CDS 101..856

FT CDS /*tag= a

FT CDS /product= "brnF"

FT CDS 853..1179

FT CDS /*tag= b

FT CDS /product= "brnE"

PN EPI096010-A1.

PD 02-MAY-2001.

PF 11-OCT-2000; 2000BP-00122057.

PR 27-OCT-1999; 99DE-01051708.

XX (DEGS) DEGUSA AG.

XX PA (KERJ) FORSCHUNGSZENTRUM JUELICH GMBH.

XX PI Kennerknecht N, Eggeling L, Sahm H, Pfeiferle W,

XX XX

DR WPI; 2001-391595/42.
 XX P-PSDB; AAB6247, AAB6248.
 PT New export genes from coryneform bacteria, useful for increasing
 PT fermentative production of branched-chain amino acids.
 XX
 PS Claim 4 (i); Page 13; 23pp; German.
 CC This invention describes a novel isolated polynucleotide (I) containing
 CC at least one sequence that (i) is 70% identical with a sequence that
 CC encodes at least one of 251 amino acid (aa) (3) or 108 aa (5)
 CC polypeptides; (ii) encodes a polypeptide at least 70% identical with (3)
 CC or (5); (iii) is the complement of (i) or (ii), or (iv) contains at least
 CC 15 consecutive bases from (i)-(iii). The invention also describes (a)
 CC coryneform microorganisms, especially Corynebacterium, transformed with
 CC one or more (i), where these are replicative DNA; (c) production of
 CC branched-chain L-aa by fermentation of coryneform bacteria in which the
 CC brnE and/or brnF genes (or equivalent sequences) are amplified,
 CC especially overexpressed; and (d) method for isolating the brnE and/or
 CC brnF genes. (I) is used for transformation of coryneform bacteria being
 CC used for fermentative production of branched-chain amino acids,
 CC specifically leucine, isoleucine and valine, which are useful in medicine
 CC and animal nutrition. (I) can also be used as source of primers and
 CC probes for isolation of related sequences. Transformation with (i)
 CC increases yield of branched-chain amino acids. This sequence encodes the
 CC Corynebacterium glutamicum ATCC 14752 brnE and brnE proteins described in
 CC the method of the invention
 XX
 SQ Sequence 1271 BP; 265 A; 328 C; 318 G; 360 T; 0 U; 0 Other;

Query Match 100.0%; Score 753; DB 4; Length 1271;
 Best Local Similarity 100.0%; Pred No. 2.5e-223;
 Matches 753; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GTGCAAAAACGCAAGATTCATTCAGGCTGAGAGTGTGCGCATCCAGAGCGCTTG 60
 Db 101 GTGCAAAAACGCAAGATTCATTCAGGCTGAGAGTGTGCGCATCCAGAGCGCTTG 160
 Qy 61 GAACCAAGATTAAGGTTATGCGCGCTAGCAAAAGCGCAAGGCTTAAGAACTCCCTT 120
 Db 161 GAACCAAGATTAAGGTTATGCGCGCTAGCAAAAGCGCAAGGCTTAAGAACTCCCTT 220
 Qy 121 GCTGAGGTTGGGCAATGATCCGATTTGATTCGCTTGTCTCTTGGTATTCAATAC 180
 Db 221 GCTGAGGTTGGGCAATGATCCGATTTGATTCGCTTGTCTCTTGGTATTCAATAC 280
 Qy 181 GGTACGAATGATGGGAGGCGCCCACTGTTTTCGCGCTGATTTTCGCGGCTCCACCGAA 240
 Db 281 GGTACGAATGATGGGAGGCGCCCACTGTTTTCGCGCTGATTTTCGCGGCTCCACCGAA 340
 Qy 241 ATGCTGATCATGCGCTGTTGAGGCGAGCGCCCTGGGCGCCATTCGCGCTCACACA 300
 Db 341 ATGCTGATCATGCGCTGTTGAGGCGAGCGCCCTGGGCGCCATTCGCGCTCACACA 400
 Qy 301 TTGCTGATGAATCTTCGCGCAGATTTCTATGCGTTTTCATTCGCGCTGATGTCGCAAA 360
 Db 401 TTGCTGATGAATCTTCGCGCAGATTTCTATGCGTTTTCATTCGCGCTGATGTCGCAAA 460
 Qy 461 AACCCCATTTGCCCGTTTCTATTCGCTTTCGCGCTTATGCAAGAGCTTACGAGTCACT 420
 Db 461 AACCCCATTTGCCCGTTTCTATTCGCTTTCGCGCTTATGCAAGAGCTTACGAGTCACT 520
 Qy 421 GCGGCGAGGCGCGAGGCTGTGCGGCGAGCTTATCTCAATGCAATAGCGTTTCAAC 480
 Db 521 GCGGCGAGGCGCGAGGCTGTGCGGCGAGCTTATCTCAATGCAATAGCGTTTCAAC 580
 Qy 481 TCCTACTGGGATTTGCGCGGCTCTCACCGAGTGGCGATCGAGAGTTGATTCCTTTGAA 540
 Db 581 TCCTACTGGGATTTGCGCGGCTCTCACCGAGTGGCGATCGAGAGTTGATTCCTTTGAA 640
 Qy 541 ATTAAGGCGCTGAGTTCGCGCTTTCGCTCTCTTTTGTACGCTGACTTGGATTCCTGC 600

Db 641 ATTAAGGCGCTCGAGTTGGCCCTTTGCTCTCTTGTGCAAGCTGACTTTGGATTCCGAC 700
 Qy 601 CGAAGCAAAAAGAGATCCCTTCTCTGCTGCTCGAGGTTTGAGCTTACCATTTCTCTT 660
 Db 701 CGAAGCAAAAAGAGATCCCTTCTCTGCTGCTCGAGGTTTGAGCTTACCATTTCTCTT 760
 Qy 661 GTGGTAATTCAGAGTCAGAGCCCTATTGCGGCGCTGATCTTCTTGGTCTGTGACC 720
 Db 761 GTGGTAATTCAGAGTCAGAGCCCTATTGCGGCGCTGATCTTCTTGGTCTGTGACC 820
 Qy 721 ATCCGTAATCTTCTTGGGAAAAGCTGCTAAA 753
 Db 821 ATCCGTAATCTTCTTGGGAAAAGCTGCTAAA 853
 RESULT 4
 AAH2112
 ID AAH2112 standard; DNA, 1271 BP.
 AC AAH2112;
 XX
 DT 05-SEP-2001 (first entry)
 DE C. glutamicum DNA encoding brnF and brnE.
 XX
 KW L-amino acid production; brnF; brnE; branched-chain amino acid;
 KW coryneform bacterium; leucine; isoleucine; valine; medicine;
 KW animal nutrition; ds.
 XX
 CS Corynebacterium glutamicum.
 XX
 FH Key Location/Qualifiers
 FT CDS 101..856
 FT /*tag= a
 FT /product= "brnF"
 FT CDS 853..1179
 FT /*tag= b
 FT /product= "brnE"
 FT
 FT
 PN EP1096010-A1.
 XX
 PD 02-MAY-2001.
 XX
 PF 11-OCT-2000; 2000EP-00122057.
 XX
 PR 27-OCT-1999; 99DE-01051708.
 XX
 PA (DEGS) DEGUSA AG.
 PA (KERJ) FORSCHUNGSZENTRUM JUELICH GMBH.
 PI Kernerkecht N, Eggeling L, Sahn H, Pfeifferle W;
 PI
 DR MPI; 2001-391595/42.
 XX
 PT New export genes from coryneform bacteria, useful for increasing
 PT fermentative production of branched-chain amino acids.
 XX
 PS Claim 4 (i); Page 17-18; 23pp; German.
 CC This invention describes a novel isolated polynucleotide (i) containing
 CC at least one sequence that (i) is 70% identical with a sequence that
 CC encodes at least one of 251 amino acid (aa) (3) or 108 aa (5)
 CC polypeptides; (ii) encodes a polypeptide at least 70% identical with (3)
 CC or (5); (iii) is the complement of (i) or (ii), or (iv) contains at least
 CC 15 consecutive bases from (i)-(iii). The invention also describes (a)
 CC protein derived from sequences of 753 bp (2) or 324 bp (4); (b)
 CC coryneform microorganisms, especially Corynebacterium, transformed with
 CC one or more (i), where these are replicative DNA; (c) production of
 CC branched-chain L-aa by fermentation of coryneform bacteria in which the
 CC brnF and/or brnE genes (or equivalent sequences) are amplified,
 CC especially overexpressed; and (d) method for isolating the brnF and/or
 CC brnE genes. (i) is used for transformation of coryneform bacteria being
 CC used for fermentative production of branched-chain amino acids,

CC Specifically leucine, isoleucine and valine, which are useful in medicine
 CC and animal nutrition. (i) can also be used as source of primers and
 CC probes for isolation of related sequences. Transformation with (i)
 CC increases yield of branched-chain amino acids. This sequence encodes the
 CC Corynebacterium glutamicum ATCC 13032 brnF and brnE proteins described in
 CC the method of the invention
 XX
 SO Sequence 1271 BP; 265 A; 328 C; 318 G; 360 T; 0 U; 0 Other;
 Query March 100.0%; Score 753; DB 4; Length 1271;
 Best Local Similarity 100.0%; Pred. No. 2.5e-223;
 Matches 753; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 GTGCAAAAAGAGAGATTCATTCAGAGCTGAGAGTGTGCGATCCAGGACGCGCTG 60
 Db 101 GTGCAAAAAGAGAGATTCATTCAGAGCTGAGAGTGTGCGATCCAGGACGCGCTG 160
 Qy 61 GAACCAATGATTAAGATTATGCGCGCTAGCAAAATGCGCAAGTCTAAACCTCCCTT 120
 Db 161 GAACCAATGATTAAGATTATGCGCGCTAGCAAAATGCGCAAGTCTAAACCTCCCTT 220
 Qy 121 GCTGCAAGTTGGGCAATGATCCGATTTGATTTGCTCTTGGTTATTCATATC 180
 Db 221 GCTGCAAGTTGGGCAATGATCCGATTTGATTTGCTCTTGGTTATTCATATC 280
 Qy 181 GGCTAGCAATGTGGGAGCGCCCACTGTTTTCGGGCTGATTTTCGGGCGTCCACCGAA 240
 Db 281 GGCTAGCAATGTGGGAGCGCCCACTGTTTTCGGGCTGATTTTCGGGCGTCCACCGAA 340
 Qy 241 ATGCTGATCATGCGCCCTGTTGTTGGGCGAGCGCCCTGAGCGCATGCGCTACACAA 300
 Db 341 ATGCTGATCATGCGCCCTGTTGTTGGGCGAGCGCCCTGAGCGCATGCGCTACACAA 400
 Qy 301 TTGCTGTGAATCTTCGCGCAAGTATTCATGCTTTTCAATCCGCTGATGTGCAAA 360
 Db 401 TTGCTGTGAATCTTCGCGCAAGTATTCATGCTTTTCAATCCGCTGATGTGCAAA 460
 Qy 361 AAOCCCATGCGCGCTTCTATTTGTTTTCGGGCTTATGAGCAAGCTACGAGCACT 420
 Db 461 AAOCCCATGCGCGCTTCTATTTGTTTTCGGGCTTATGAGCAAGCTACGAGCACT 520
 Qy 421 GCGGCAAGGCGCGAGCTGTGCGGCGAGCGAATTATCTGAATGCAAAATAGCGTTTCA 480
 Db 521 GCGGCAAGGCGCGAGCTGTGCGGCGAGCGAATTATCTGAATGCAAAATAGCGTTTCA 580
 Qy 481 TCCACTGAGTATTCGGCGGCTCTACCGGAGTGGCGATGCAAGTTGATTCCTTTGAA 540
 Db 581 TCCACTGAGTATTCGGCGGCTCTACCGGAGTGGCGATGCAAGTTGATTCCTTTGAA 640
 Qy 541 ATTAAGGCGCTCGAGTTGCGCTTGTCTCTCTTTGTGCAAGCTGACTTTGATTCCTG 600
 Db 641 ATTAAGGCGCTCGAGTTGCGCTTGTCTCTCTTTGTGCAAGCTGACTTTGATTCCTG 700
 Qy 601 CGAAGCAAAAAGAGATCCCTTCTCTGCTGCTCGAGGTTTGAGCTTACCATTTCTCTT 660
 Db 701 CGAAGCAAAAAGAGATCCCTTCTCTGCTGCTCGAGGTTTGAGCTTACCATTTCTCTT 760
 Qy 661 GTGGTAATTCAGAGTCAGAGCCCTATTGCGGCGCTGATCTTCTTGGTCTGTGACC 720
 Db 761 GTGGTAATTCAGAGTCAGAGCCCTATTGCGGCGCTGATCTTCTTGGTCTGTGACC 820
 Qy 721 ATCCGTAATCTTCTTGGGAAAAGCTGCTAAA 753
 Db 821 ATCCGTAATCTTCTTGGGAAAAGCTGCTAAA 853
 RESULT 5
 AAH64966
 ID AAH64966 standard; DNA, 349980 BP.
 AC AAH64966;
 XX
 DT 26-SEP-2001 (first entry)

XX Sequence 564 BP; 91 A; 167 C; 144 G; 162 T; 0 U; 0 Other;

Query Match 73.4%; Score 553; DB 7; Length 564;
Best Local Similarity 99.8%; Pred. No. 2.7e-161;
Matches 564; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

```

QY 189 ATGGTGGAGAGCCCACTGTTTCGGGCGTGAATTTTCGGGGCTCCACGAAATGCTGT 248
DB 1 ATGGTGGAGAGCCCACTGTTTCGGGCGTGAATTTTCGGGGCTCCACGAAATGCTGT 60
QY 249 CATGCGCTCGTGTGTGGGGCGAGCGCCCTGGGCGGCATCGGCTCACCACATTTGCTGT 308
DB 61 CATGCGCTCGTGTGTGGGGCGCA-CGCCCTGGGGCGGCATCGGCTCACCACATTTGCTGT 119
QY 309 GAACCTCCCGCAGCTATTTCTAGTCCGTTTCATTTCCCGCGCATGTGTGTAAAAACCCAT 368
DB 120 GAACCTCCCGCAGCTATTTCTAGTCCGTTTCATTTCCCGCGCATGTGTGTAAAAACCCAT 179
QY 369 TGCCCGTTTCTATTCGATTTTCGGGCTTATCGAGAGGCTTACGAGTCACTGCGGCGAG 428
DB 180 TGCCCGTTTCTATTCGATTTTCGGGCTTATCGAGAGGCTTACGAGTCACTGCGGCGAG 239
QY 429 GCCCGCAGGCTGTGGCGGCTGGGCACTTATCTGAATGCAATAGGCTTCACTCTACTG 488
DB 240 GCCCGCAGGCTGTGGCGGCTGGGCACTTATCTGAATGCAATAGGCTTCACTCTACTG 299
QY 489 GGTATTGCGGCGTCTACCGGAGTGGCGATCGCAGAGTGTATCTTTGAAATTAAGG 548
DB 300 GGTATTGCGGCGTCTACCGGAGTGGCGATCGCAGAGTGTATCTTTGAAATTAAGG 359
QY 549 CTTGAGTTCGCCCTTTGCTCTCTTTGTACGCTGACTTTGATTTCTGCGGAAAGAA 608
DB 360 CTTGAGTTCGCCCTTTGCTCTCTTTGTACGCTGACTTTGATTTCTGCGGAAAGAA 419
QY 609 AAGGAGATGCTCTCTCTGCTGCTCGAGAGTTGAGTTGACATGCTGCTTGTGTAT 668
DB 420 AAGGAGATGCTCTCTCTGCTGCTCGAGAGTTGAGTTGACATGCTGCTTGTGTAT 479
QY 669 TCCAGGTACAGCCCTATTTTGGGCGCTGCTGATCTTTTGGGTCTGTGACATCGGTA 728
DB 480 TCCAGGTACAGCCCTATTTTGGGCGCTGCTGATCTTTTGGGTCTGTGACATCGGTA 539
QY 729 CTCTCTTGTGGGAAAGGCTGCTAA 753
DB 540 CTCTCTTGTGGGAAAGGCTGCTAA 564

```

RESULT 7
AAF61693/c
ID AAF61693 standard; DNA; 397 BP.

```

XX AAF61693;
AC
XX 12-JUL-2001 (first entry)
DT
XX
DE C. glutamicum lrp encoding DNA fragment SEQ ID 7.
XX
KW Fermentation; L-amino acid production; lrp gene; lysine; isoleucine;
KW medicine; animal feed supplement; ds.
XX
OS Corynebacterium glutamicum.
XX
PN EP1090993-A1.
XX
PD 11-APR-2001.
XX
PF 29-SEP-2000; 2000EP-00121159.
XX
PR 05-OCT-1999; 99DE-01047792.
XX
PA (DEGSA ) DEGUSA-HUELS AG.
XX

```

PI Moekkel B, Pfeifferle W, Puehler A, Kalinowski J, Bathe B;
XX WPI; 2001-29927/31.
DR

XX New lrp gene from coryneform bacteria, used to prepare transformants with
XX increased synthesis of amino acids, particularly lysine and isoleucine.
XX
XX Example 3; Page 16; 22pp; German.

XX This invention describes a novel isolated nucleic acid (I) from
XX coryneform bacteria used for the fermentative production of selected L-
XX amino acids, by fermenting the amino acid-producing coryneform in which
XX at least the lrp gene has been weakened or amplified, then isolating
XX amino acids that have accumulated in the medium or cells. (I) is used to
XX transform coryneforms for production of L-amino acids, specifically
XX lysine and isoleucine, which are used in medicine and particularly as
XX animal feed supplement. It may also be used as probes and primers for
XX production of amino acids, especially of L-lysine. This sequence encodes
XX a fragment of the Corynebacterium glutamicum lrp protein which is used in
XX the method described in the invention

SEQ Sequence 397 BP; 96 A; 100 C; 104 G; 97 T; 0 U; 0 Other;

Query Match 36.3%; Score 273.6; DB 4; Length 397;
Best Local Similarity 98.6%; Pred. No. 1.7e-74;
Matches 276; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```

QY 1 GTGCAAAAAGCAGAGATTCATTCAAGCTGAGAGTGTGCGATCAAGAGAGCCCTG 60
DB 284 GTGCAAAAAGCAGAGATTCATTCAAGCTGAGAGTGTGCGATCAAGAGAGCCCTG 225
QY 61 GAACCAAGTATTAAGTTATCGGCGCTAGAAATCGCGAAGTCTAAAAACCTCCCTT 120
DB 224 GAACCAAGTATTAAGTTATCGGCGCTAGAAATCGCGAAGTCTAAAAACCTCCCTT 165
QY 121 GCTGAGGTTTGGGCACTGACCGCATTTGATTTGCTGTTGCTGTTATTAATAC 180
DB 164 GCTGAGGTTTGGGCACTGACCGCATTTGATTTGCTGTTGCTGTTATTAATAC 105
QY 181 GCTACGAATGTGGGAGACCCCACTGTTTTCGAGCTGATTTTGGCGGCTCCACGAA 240
DB 104 GCTACGAATGTGGGAGACCCCACTGTTTTCGAGCTGATTTTGGCGGCTCCACGAA 45
QY 241 ATGCTGATCATGCGCCCTGTTGTGGGGCGAGCGGCCCTGG 280
DB 44 ATGCTGATCATGCGCCCTGTTGTGGGGCGAGCGGCCCGGG 5

```

RESULT 8
AAF61695/c
ID AAF61695 standard; DNA; 778 BP.

```

XX AAF61695;
AC
XX 12-JUL-2001 (first entry)
DT
XX
DE C. glutamicum lrp encoding DNA fragment SEQ ID 9.
XX
KW Fermentation; L-amino acid production; lrp gene; lysine; isoleucine;
KW medicine; animal feed supplement; ds.
XX
OS Corynebacterium glutamicum.
XX
PN EP1090993-A1.
XX
PD 11-APR-2001.
XX
PF 29-SEP-2000; 2000EP-00121159.
XX
PR 05-OCT-1999; 99DE-01047792.
XX
PA (DEGSA ) DEGUSA-HUELS AG
XX

```


CC coryneform bacterium. Coryneform bacteria are useful for producing amino
 CC acids, nucleic acids, vitamins, saccharides and organic acids,
 CC particularly L-lysine. The present sequence is a nucleic acid described
 CC in the exemplification of the invention. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from the European Patent Office

XX SQ Sequence 309400 BP; 70133 A; 86477 C; 83115 G; 69675 T; 0 U; 0 Other;

Query Match 16.9%; Score 127.4; DB 5; Length 309400;
 Best Local Similarity 51.5%; Pred. No. 9.6e-28;
 Matches 318; Conservative 0; Mismatches 296; Indels 3; Gaps 1;

QY 122 CTGACAGTTTGGGACATGACCCGATGTTGATTCGTTGCTCTTGTGTTATTCATACG 181
 DB 289669 CGGTCCGTTTGGGGCGATTCCTCGGTTTGCCCTTGGCGTTGATGATCGACAG 289910
 QY 182 GCTACGAATGATGGGAGCCCACTGTTTCCGGCTGATTTTCCGGGCTCCACCGAAA 241
 DB 289909 GTTTCGCTGTGTGTGAGACCCGATTTCTCTTCTGATCTATGCGGTTGATGAA 289850
 QY 242 TGTGTGATCGCCCTGTGTGTGTGGGAGAGCCCTTGGGCGCATCGCTTACCAAT 301
 DB 289849 TTCTGGGAATCGGCATGTGTACCGAGTATCGGCCCGTTTTCGGGCGGTCTGTG 289790
 QY 302 TGTGTGGAATCCGCGACGATTTCTATGCGTTTCACTCCCGCTGCAATGTGTGAAA 361
 DB 289789 TCAATGTAATTTCCGCAATTTTCTACGCTCTCACCTTCCAGCCACCGCATAA 289730
 QY 362 ACCCAATGCGCGTTTCTATTCGTTTTCGCGCTTATCGAAGGCTTACGAGTCA 421
 DB 289729 CCGGCGCGCGCGCGCTTATTCACCTAGCGGCTTACCAAGAGTCTTACGCAAT 289670
 QY 422 CGGCGAGCGCGAGGTGTGTGGCGG---GGACTTATCTCAATGCAAAATGAGCTT 478
 DB 289669 CACCGCGCCACCTGGGATATAGTGGACGCGGCGCTTACCGTTCAAATTTGTGCC 289610
 QY 479 ACTCTACTAGGTTATCGCGGTCTCACCGAGTGGCGATCGAGAGTTGATTCCTTTG 538
 DB 289609 AAGCTGTGTGTATCCAGAAATATATGGCGCTTGTGTGTAAGTGTGCGCGAG 289550
 QY 539 AAATTAAGGCGCTGAGTTGCGCTTGTCTCTCTTGTGACAGCTGATTTGATTCCT 598
 DB 289549 ATCTAAAGGCAATGATTTTGCCTGACCGCGCTGTTGTGTGTGCGGTGGAG 289490
 QY 599 GCCGAAGAAAGAGATCCCTTCTGCTGTGCGAGTTTGAAGTTTCAACATTGCTC 658
 DB 289489 TCAAAATTAACAGATATATCGCTGATTTCCGCGATGTATGCGCTGTGTTCCG 289430
 QY 659 TTGTGTAATTCAGGTGAGCCCTATTTGGCGCGCTGTGATCTTTGGGTCTGTGA 718
 DB 289429 GTTTTGTGGCGCGCGAGCAATGCTGTGTATCGTTTGAACAGTACTTTTGTAT 289370
 QY 719 CCAATCGGTAATCTTC 735
 DB 289369 TTCTCCGCTCCGCTTC 289353

RESULT 12
 ACA01968
 ID ACA01968 standard; DNA; 177 BP.
 XX
 AC ACA01968;
 XX
 DT 04-JUN-2003 (first entry)
 XX
 DE C. glutamicum derived ORF SEQ ID 1959.
 XX
 KW Coryneform; nucleic acid array; fermentation; culture; ds.
 XX
 OS Corynebacterium glutamicum.
 XX
 PN DB10128510-A1.

XX 19-DEC-2002.
 PD
 XX 13-JUN-2001; 2001DE-01028510.
 XX
 PF 13-JUN-2001; 2001DE-01028510.
 XX
 PR 13-JUN-2001; 2001DE-01028510.
 XX
 PA (DEGS) DEGUSA AG.
 XX
 PI Farwick M, Moeckel B, Pfeifferle W, Bathe B, Hutmacher K;
 XX
 DR WPI; 2003-279970/28.
 XX
 PT New nucleic acid array useful for monitoring mRNA expression of
 PT Corynebacterium glutamicum during fermentation, comprising nucleic acid
 PT from Corynebacterium glutamicum.

PS Claim 1; Page 647; 709pp; German.
 XX
 CC This invention describes a novel nucleic acid array involving
 CC Corynebacterium glutamicum polynucleotides. The arrays are used to
 CC analyse C. glutamicum, particularly for monitoring a fermentation process
 CC to determine expression levels of C. glutamicum cellular mRNA. Such
 CC monitoring particularly differentiates between expression levels of
 CC different strains of C. glutamicum and allows the adjustment of different
 CC culture and fermentation conditions. ACA00010-ACA02188 represent C.
 CC glutamicum derived polynucleotides described in the disclosure of the
 CC invention

XX SQ Sequence 177 BP; 54 A; 40 C; 47 G; 36 T; 0 U; 0 Other;

Query Match 14.1%; Score 106; DB 7; Length 177;
 Best Local Similarity 100.0%; Pred. No. 1.5e-22;
 Matches 106; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTGCAAAAAGCAAGATTCATTCAAGCTGAGGTGTGCCATCCAGGCGCCCTG 60
 DB 72 GTGCAAAAAGCAAGATTCATTCAAGCTGAGGTGTGCCATCCAGGCGCCCTG 131
 QY 61 GAACAGATGATTAAGTTATCGGCGCTACGAAATCGCGCAAGTTC 106
 DB 132 GAACAGATGATTAAGTTATCGGCGCTACGAAATCGCGCAAGTTC 177

RESULT 13

AAF61688/C
 ID AAF61688 standard; DNA; 715 BP.

XX AAF61688;
 AC
 XX
 DT 12-JUN-2001 (first entry)
 XX
 DE C. glutamicum lrp encoding DNA.
 XX
 KW Fermentation; L-amino acid production; lrp gene; lysine; isoleucine;
 KW medicine; animal feed supplement; ds.
 XX
 OS Corynebacterium glutamicum.
 XX
 OS

XX Key Location/Qualifiers
 XX FH -35_signal 62..67
 XX FT -10_signal 88..93
 XX FT CDS 151..615
 XX FT /*tag= b
 XX FT /*tag= c
 XX FT /product= "lrp"

XX EP1090993-A1.
 XX
 PD 11-APR-2001.
 XX
 XX 29-SEP-2000; 2000EP-00121159.

XX	PR	05-OCT-1999;	99DE-01047792.
XX	XX	(DEGS) DEGUSSA-HUELS AG.	
PI	PI	Moeckel B, Pfeifferle W, Puenher A, Kalinowski J, Bathe B;	
DR	DR	WPI; 2001-292927/31.	
XX	P-PSDB;	AAB70881.	
PT	New lrp gene from coryneform bacteria, used to prepare transformants with		
PT	increased synthesis of amino acid, particularly lysine and isoleucine.		
PS	Claim 4; Page 14-15; 22pp; German.		
XX	XX	This invention describes a novel isolated nucleic acid (I) from	
CC	CC	coryneform bacteria used for the fermentative production of selected L-	
CC	CC	amino acids, by fermenting the amino acid-producing coryneform in which	
CC	CC	at least the lrp gene has been weakened or amplified, then isolating	
CC	CC	amino acids that have accumulated in the medium or cells. (I) is used to	
CC	CC	transform coryneforms for production of L-amino acids, specifically	
CC	CC	lysine and isoleucine. Which are used in medicine and particularly as	
CC	CC	animal feed supplement. It may also be used as probes and primers for	
CC	CC	isolating related sequences. Regulating expression of (I) improves	
CC	CC	production of amino acids, especially of L-lysine. This sequence encodes	
CC	CC	the Corynebacterium glutamicum lrp protein which is used in the method	
CC	CC	described in the invention	
SQ	SQ	Sequence 715 BP; 166 A; 173 C; 178 G; 198 T; 0 U; 0 Other;	
OY	Query Match	11.3%; Score 85; DB 4; Length 715;	
Dd	Best Local Similarity	100.0%; Pred. No. 9.6e-16;	
OY	Matches 85; Conservative	0; Mismatches 0; Indels 0; Gaps 0	
OY	1 GTGCAAAAAGCGAGAAGATTCATTCAACGCCTGGAGGTGCCATCCAAGGACCCTG 60		
Dd	85 GTGCAAAAAACGAAGAGATTCATTCAACCTCGAGCTGTGCCATCCAAGGACCCTG 26		
OY	61 GAACCATGATTAAGGTTATCGGC 85		
Dd	25 GAACCATGATTAAGGTTATCGGC 1		
XX	RESULT 14		
AD	ADA71938/c		
ID	ADA71938 standard; DNA; 2000 BP.		
XX	ADA71938;		
AC	20-NOV-2003 (first entry)		
DT	Rice gene, SEQ ID 5263.		
XX	DE		
XX	Plant; bacterial infection; fungal infection; viral infection; rice;		
KM	gene; ds.		
XX	Oryza sativa.		
OS	WO2003000898-A1.		
XX	PN		
XX	PD		
XX	03-JAN-2003.		
PF	22-JUN-2001; 2001WO-IB001105.		
XX	XX		
PR	22-JUN-2001; 2001WO-IB001105.		
XX	PA		
XX	(SYGN) SYNGENTA PARTICIPATIONS AG.		
PI	Chang H, Chen W, Cooper B, Glazebrook J, Gott SA, Hou Y;		
PI	Kategirai F, Qian S, Tao Y, Whitnam S, Xie Z, Zhu T, Zou G;		
XX	WPI; 2003-175290/17.		
DR			

Query Match	Best Local Similarity	6.2%	Score 47	DB 7	Length 2000
Matches	75	Conservative 328	Mismatches 323	Indels 11	Gaps 2
Qy	8	AAACGACAGAGATTGATTCAGACCTCGAGAGTGTGCGCATCAAGACGAGCCCTGGAACGAG	67		
Db	736	RMMSAMRYCSRMKCACTKYASASARWTKRARSRRRRMYRRKGGTYRYRWRSCMT	677		
Qy	68	ATGATAAGGTTATGGGGCGGTACGAAATCGCGCAAGGCTTAAAACTCCCTGCTGCGAG	127		
Db	676	RAAMSRKRKMAAGSMKSMCMWTRGRASNMYSKYSCKACKKTRWTSWMSGTGMS	617		
Qy	128	GTTTGGGACATGACCCGATGGTATGTGGTTGGCTCTTGTATTCAATACGGCTACG	187		
Db	616	S--YKSMWTSKMSYMKMTCTMYTSMKSTRSKMRWMSGMSRMVRRMKRKRKYMR	559		
Qy	188	AATGGTGGGACGCCCACTGTTTCCGGCCGATTTTGGGGGGTCCACCGAAATGCTGG	247		
Db	558	YMKMKCTMRMCNCRWGTMTTSSRMWTRGRKARATSKRYRMYYKRYKTYIYGM	499		
Qy	248	TCATCGCCCT-----CGTTGGGCGCAGGCGCCCTGGGCGCATCGCCTACCA	298		
Db	498	YMKSYMMRYGYCKACKCCCYAMCKAKAAYSMMMYWRYKTSKMMRBSTKYSMMWYKCR	439		
Qy	299	CATTGCGTGAACCTTCGCGCACGATTCATGAGGTTTCACTCCGCTGCATGTGATCA	358		
Db	438	SMKRGAGCYGCKMWTGYSYGMKYTTMSYKTSBKCYKRYMYTGMWYMYTAYSS	379		
Qy	359	AAACCCCATTTGCCGGTTCTATTGCGTTTTGCGGCTTATCGAAGACCTACGACATCA	418		
Db	378	MMTWYYVYAKWYKWTYKRGTSWYKGYKXKRYCTWVCYMKMRCYMRDMRKRTYKS	319		
Qy	419	CTGGGCGCAGCGCCGAGGCTGGTGGGGGTGGCGACCTATCTCAATCAATAGCGTTTC	478		
Db	318	RCYCWRYATYMCCTCRKRGWTSRSRSMRTAGKMKMSMSWMSKCSYSWYMYTKMKKEY	259		
Qy	479	ACTCTACTAGGTATTGGCGGCTCTCACCGGATGGCGATCGCAGAATTATTCCTTTTG	538		
Db	258	YMSGNAARSAGTSSRSARATRYKYSTSRBAMRBACGMSACRRVSRSTYCGGSCGS	199		
Qy	539	AAATTAAGGCGCTCGAGTTGCGCCCTTTGCTCTCTCTTTGACGCTACATTGGATTCT	598		
Db	198	SKTKWKSRSKMTTSSWCSGCTCYCYAGACMCSCMSMMYMSGCGYTGMRKRSYSM	139		
Qy	599	GCCGAAAGAAAGACGATCCTTCTCTGCTGCTGCGCAGGTTGAGCTTACCATGCTC	658		
Db	138	CCKRYCSGCTKYCSYGYRYCKWYKMYKYCYCYWYMSYMRMYMKCMCSGCSNM	79		
Qy	659	TTGNGTATTCAGGCTCAGGCCCATTTGCGGCGCTCATATCTTGTGGTGTGTTGA	718		
Db	78	SCAYCSTSSSRMSMYTAAKMGCGSSGMYRMSKCSWYSKYSCKYTSKCTKRYK	19		
Qy	719	CCATCGGATCTTCTTC	735		

Db 18 YCWSGYSVMCTSGTY 2

RESULT 15
ID AAA10594/c
AAA10594 standard, DNA, 10732 BP.

AC AAA10594;

DT 29-JUN-2000 (first entry)

DE Gene encoding a subunit of cellulose synthase.

KW Cellulose synthase; cellulose production; increase yield; ds.

OS Vigna angularis.

XX JP2000060568-A.

XX 29-FEB-2000.

XX 26-AUG-1998; 98JP-00239998.

XX 26-AUG-1998; 98JP-00239998.

XX (MIZU/) MIZUNO K.

XX (OJIP) OJI PAPER CO.

XX WPI; 2000-342371/30.

XX P-PSDB; AAY85179.

PT A gene encoding a cellulose synthetic equipment - for the improvement in

XX the amount of cellulose synthesized in a plant body.

XX Claim 2; Page 14-21; 32pp; Japanese.

XX This sequence represents a gene encoding a subunit of the cellulose

CC synthase complex of Vigna angularis. The invention relates to subunits of

CC cellulose synthetic equipment, that can be used to increase the amount of

CC cellulose synthesized by a plant. The proteins and genes encoding them

CC can also be used to improve the properties of the cellulose being

CC produced by a plant

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

Query Match 5.5%; Score 41.2; DB 3; Length 10732;

Best Local Similarity 14.7%; Pred. No. 0.13; Mismatches 189; Indels 1; Gaps 1;

Matches 64; Conservative 182; Mismatches 189; Indels 1; Gaps 1;

262 GTGAGGCGAGCGCCCTGAGCGCGCATGCGCTCACACCATTCGTTGTAATTCGCGCAC 321

9949 RYSSSTGNCYSYCTTSTTNRCTTTSYNSSTTBTBYSYSTSRGYSRGSDS 9890

322 GTATTCTATGCGTTTCATTCGCGCTGATGTCATGTCATGTCATGTCATGTCATGTCAT 381

9889 RGNCTNSTNCYDASTDSTBYSRCYTSYSTDSTSTSTSTSTSTSTSTSTSTSTSTSTST 9830

382 TCGGTTTTCGCGCTTATCGACGAGCGTACGAGCTGAGCGCGCGCGCGCGCGCGCGCG 441

9829 SDSTSTYRCRSRSDYDASTDSTBYSRCYTSYSTDSTSTSTSTSTSTSTSTSTSTSTST 9770

442 TCGGCGTGGCGACTTATCGATGCAATGCAATGCGTTTCTACTGCGTATTCGCGGT 501

9769 GYDANSTSRYSSTYSSTYSSTYSSTYSSTYSSTYSSTYSSTYSSTYSSTYSSTYS 9710

502 CTCACCGAGTGGCGATGCGAGAGTTCCTTTGAAATTAAGGCGCTCG-AGTTGCG 560

9709 YTBYSRRCRCYDASTDSTBYSRCYTSYSTDSTSTSTSTSTSTSTSTSTSTSTSTSTST 9650

561 CCTTGCTCTCTCTTGTGACGCTGACTTGAATTCCTGCGGACGAAAGCAAGATCC 620

9649 TTSRCTBYSBTBTBTRCAKCTBDSTAKNSTSTSTSTSTSTSTSTSTSTSTSTSTSTST 9590

QY 621 TTCTGCTGCTGCGAGGTTGAGCTTCACCATTCGCTCTTGTTGTAATTCAGGTCAGGC 680

Db 9589 CYSCTDSDSTCYSTTSTYAYSCCTTSRGSYDASTSTSTYRCCTTSTYSTTSDC 9530

QY 681 CCTATTGCGCGCGCTG 696

Db 9529 SDYSTTBTNSTYSSDS 9514

Search completed: Apr-11 19, 2004, 11:42:00
Job time : 352.669 secs

Tue Apr 20 06:47:16 2004

us-10-608-504-2.rnt

Page 1

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

CM nucleic - nucleic search, using sw model

Run on: April 19, 2004, 10:57:33 ; Search time 63.4023 Seconds
(without alignments)
6590.898 Million cell updates/sec

Title: US-10-608-504-2

Perfect score: 1 gtgcacaaacgcaagagat.....tctcggaagagctcctaaa 753

Sequence: 1 gtgcacaaacgcaagagat.....tctcggaagagctcctaaa 753

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 682709 segs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:
1: /cgn2_6/pdata/2/ina/5A.COMB.seq:*
2: /cgn2_6/pdata/2/ina/5B.COMB.seq:*
3: /cgn2_6/pdata/2/ina/6A.COMB.seq:*
4: /cgn2_6/pdata/2/ina/6B.COMB.seq:*
5: /cgn2_6/pdata/2/ina/PCUS.COMB.seq:*
6: /cgn2_6/pdata/2/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	753	100.0	753	US-09-471-803A-2	Sequence 2, Appli
2	753	100.0	1271	US-09-471-803A-1	Sequence 1, Appli
3	753	100.0	1271	US-09-471-803A-6	Sequence 6, Appli
4	58.6	7.8	756	US-09-489-039A-3016	Sequence 14, Appli
5	46.2	6.1	7218	US-08-232-463-14	Sequence 1266, Ap
6	39.8	5.3	768	US-09-107-532A-1266	Sequence 11376, A
7	39.2	5.2	1038	US-09-252-991A-11376	Sequence 11418, A
8	39.2	5.2	1086	US-09-252-991A-11418	Sequence 8976, Ap
9	39.2	5.2	1399	US-09-621-976-8976	Sequence 3205, Ap
10	36.8	4.9	1140	US-09-252-991A-2761	Sequence 2761, Ap
11	36.8	4.9	2061	US-09-020-956-12	Sequence 12, Appli
12	36.4	4.8	751	US-09-030-607-12	Sequence 12, Appli
13	36.4	4.8	751	US-09-439-313-12	Sequence 12, Appli
14	36.4	4.8	751	US-09-352-616A-12	Sequence 12, Appli
15	36.4	4.8	751	US-09-232-149A-12	Sequence 12, Appli
16	36.4	4.8	751	US-09-159-812-12	Sequence 12, Appli
17	36.4	4.8	751	US-09-636-215-12	Sequence 12, Appli
18	36.4	4.8	751	US-09-685-166A-12	Sequence 12, Appli
19	36.4	4.8	751	US-09-115-453-12	Sequence 12, Appli
20	36.4	4.8	751	US-09-688-489-12	Sequence 1666, A
21	36.4	4.8	430	US-09-621-976-1666	Sequence 18033, A
22	35.4	4.7	474	US-09-621-976-18033	Sequence 2399, A
23	35.4	4.7	801	US-09-328-352-2399	Sequence 2, Appli
24	35.2	4.7	4403765	US-09-103-840A-2	Sequence 1, Appli
25	35.2	4.7	4411529	US-09-103-840A-1	Sequence 5047, Ap
26	34.8	4.6	867	US-09-489-039A-5047	

28	34.6	4.6	666	2	US-08-875-034A-1	Sequence 1, Appli
29	34.4	4.6	744	4	US-09-489-039A-1307	Sequence 1307, Ap
30	34.4	4.6	1224	4	US-09-266-965-22	Sequence 22, Appli
31	34.4	4.6	12249	4	US-09-266-965-74	Sequence 74, Appli
32	34.4	4.6	18311	4	US-09-266-965-96	Sequence 96, Appli
33	33.4	4.4	1875	3	US-08-422-869A-21	Sequence 21, Appli
34	33.4	4.4	2949	3	US-08-433-522A-1	Sequence 1, Appli
35	33.4	4.4	2949	3	US-09-135-166-1	Sequence 1, Appli
36	33.4	4.4	2949	3	US-08-942-046-1	Sequence 1, Appli
37	33.4	4.4	2950	3	US-08-433-522A-5	Sequence 5, Appli
38	33.4	4.4	2950	3	US-09-135-166-5	Sequence 5, Appli
39	33.4	4.4	2950	3	US-08-942-046-5	Sequence 5, Appli
40	33.4	4.4	2984	3	US-08-433-522A-3	Sequence 3, Appli
41	33.4	4.4	2984	3	US-09-135-166-3	Sequence 3, Appli
42	33.4	4.4	2984	3	US-08-942-046-3	Sequence 3, Appli
43	33.4	4.4	2987	3	US-08-433-522A-55	Sequence 55, Appli
44	33.4	4.4	2987	3	US-09-135-166-55	Sequence 55, Appli
45	33.4	4.4	2987	3	US-08-942-046-55	Sequence 55, Appli

ALIGNMENTS

RESULT 1	US-09-471-803A-2	Application US/09471803A
Sequence 2, Appli	US-09-471-803A-2	
Patent No. 6613545		
GENERAL INFORMATION:		
APPLICANT: KENNERKNECHT, NICOLE		
APPLICANT: SAM, HERMANN		
APPLICANT: EGGELING, LOTHAR		
APPLICANT: PFEFFERLE, WALTER		
TITLE OF INVENTION: NUCLEOTIDE SEQUENCE CODING FOR THE EXPORT OF		
TITLE OF INVENTION: BRANCHED CHAIN AMINO ACIDS, PROCESS FOR THE		
TITLE OF INVENTION: ISOLATION THEREOF AND USE THEREOF		
FILE REFERENCE: 21123/265496/MAS		
CURRENT APPLICATION NUMBER: US/09/471.803A		
PRIOR FILING DATE: 1999-12-23		
PRIOR APPLICATION NUMBER: DE 199 51 708.8		
PRIOR FILING DATE: 1999-10-27		
NUMBER OF SEQ ID NOS: 12		
SOFTWARE: Patentin Ver. 2.1		
SEQ ID NO 2		
LENGTH: 753		
TYPE: DNA		
ORGANISM: Corynebacterium glutamicum		
FEATURE:		
NAME/KEY: CDS		
LOCATION: (1) ..(753)		
OTHER INFORMATION: bnf		
OTHER INFORMATION: ATCC14752		
US-09-471-803A-2		
Query Match	100.0%; Score 753; DB 4; Length 753;	
Best Local Similarity	100.0%; Pred. No. 2.2e-231;	
Matches	753; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1 GTGCAAAAACGCAAGATTCATTCAAGCTGAGGTGTCGCATTCAGGACGACCTG 60	
DB	1 GTGCAAAAACGCAAGATTCATTCAAGCTGAGGTGTCGCATTCAGGACGACCTG 60	
QY	61 GAACCAATGATTAAGTTATCGCGCTACGAAATCGCGAAGGTCTAAACCTCCCTT 120	
DB	61 GAACCAATGATTAAGTTATCGCGCTACGAAATCGCGAAGGTCTAAACCTCCCTT 120	
QY	121 GCTGAGGTTGGGACATGACCCGATTTGCTTTGCTCTCTTTGTTATTCATAC 180	
DB	121 GCTGAGGTTGGGACATGACCCGATTTGCTTTGCTCTCTTTGTTATTCATAC 180	
QY	181 GGCTACGATGATGAGGACGACCCGCTGTTTCCGGCTGATTTCCGGGCTCCACCGAA 240	
DB	181 GGCTACGATGATGAGGACGACCCGCTGTTTCCGGCTGATTTCCGGGCTCCACCGAA 240	

QY 241 ATGCTGTCATCGCCCTGTTGTGGGCGCAGCCGCCCTGAGCCATCGGCTCAACCA 300
DB 241 ATGCTGTCATCGCCCTGTTGTGGGCGCAGCCGCCCTGAGCCATCGGCTCAACCA 300
QY 301 TTGCTGTCATCGCCCTGTTGTGGGCGCAGCCGCCCTGAGCCATCGGCTCAACCA 360
DB 301 TTGCTGTCATCGCCCTGTTGTGGGCGCAGCCGCCCTGAGCCATCGGCTCAACCA 360
QY 361 AACCCATTCGCCCTGTTGTGGGCGCAGCCGCCCTGAGCCATCGGCTCAACCA 420
DB 361 AACCCATTCGCCCTGTTGTGGGCGCAGCCGCCCTGAGCCATCGGCTCAACCA 420
QY 421 GCGGCGAGCCCGCAGGCTGTGCGGCGTGGCACTTATCGAAGCCATCGAGTCACT 480
DB 421 GCGGCGAGCCCGCAGGCTGTGCGGCGTGGCACTTATCGAAGCCATCGAGTCACT 480
QY 481 TCTCTAGTATTCGCGCGCTCTCAACCGAGTGGCAATCGAGTTCATTCCTTTGAA 540
DB 481 TCTCTAGTATTCGCGCGCTCTCAACCGAGTGGCAATCGAGTTCATTCCTTTGAA 540
QY 541 ATTAAGGCTTCGAGTTCGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 600
DB 541 ATTAAGGCTTCGAGTTCGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 600
QY 601 CGAAGCAAAAGAGATCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 660
DB 601 CGAAGCAAAAGAGATCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 660
QY 661 GTGTATTCAGGTCAGGCTTCGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 720
DB 661 GTGTATTCAGGTCAGGCTTCGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 720
QY 721 ATCCGTAATTCCT 753
DB 721 ATCCGTAATTCCT 753

RESULT 2

US-09-471-803A-1
Sequence 1, Application US/09471803A

Patent No. 6613545
GENERAL INFORMATION:
APPLICANT: KENNERNKUCHT, NICOLE
APPLICANT: SAHM, HERMANN
APPLICANT: PFEFFERLE, WALTER
TITLE OF INVENTION: NUCLEOTIDE SEQUENCE CODING FOR THE EXPORT OF
TITLE OF INVENTION: BRANCHED CHAIN AMINO ACIDS, PROCESS FOR THE
FILE REFERENCE: 21123/265496/MAS
CURRENT APPLICATION NUMBER: US/09/471, 803A
PRIOR FILING DATE: 1999-12-23
PRIOR APPLICATION NUMBER: DE 199 51 708.8
NUMBER OF SEQ ID NOS: 12
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1
LENGTH: 1271
TYPE: DNA
ORGANISM: Corynebacterium glutamicum
FEATURE:
NAME/KEY: gene
LOCATION: (101)..(853)
OTHER INFORMATION: DIME
NAME/KEY: gene
LOCATION: (853)..(1176)
OTHER INFORMATION: DIME
OTHER INFORMATION: ATCC14752
US-09-471-803A-1

Query Match 100.0%; Score 753; DB 4; Length 1271;
Best Local Similarity 100.0%; Pred. No. 2,9e-21;
Matches 753; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTGCAAAAAGCAGAGATTCATTCAGGCTGAGGTGTGCGCATCCAGGAGCCCTG 60
DB 101 GTGCAAAAAGCAGAGATTCATTCAGGCTGAGGTGTGCGCATCCAGGAGCCCTG 160
QY 61 GAACAGATGATAAGGTTATCGGCTTACGAAATCGGCAAGGTTAAAGCTCCCTT 120
DB 161 GAACAGATGATAAGGTTATCGGCTTACGAAATCGGCAAGGTTAAAGCTCCCTT 220
QY 121 GCTGAGGTTGGGCTATACCCGATTTGGTTATTTGGTTATTTGGTTATTTGAATAC 180
DB 221 GCTGAGGTTGGGCTATACCCGATTTGGTTATTTGGTTATTTGGTTATTTGAATAC 280
QY 181 GGTACGAATGTGGGAGCCCGCATGTTTTCGCGCTGATTTTCGCGGCTCCACCGAA 240
DB 281 GGTACGAATGTGGGAGCCCGCATGTTTTCGCGCTGATTTTCGCGGCTCCACCGAA 340
QY 241 ATGCTGTCATCGCCCTGTTGTGGGCGCAGCCGCCCTGAGCCATCGGCTCAACCA 300
DB 341 ATGCTGTCATCGCCCTGTTGTGGGCGCAGCCGCCCTGAGCCATCGGCTCAACCA 400
QY 301 TTGCTGTCATCGCCCTGTTGTGGGCGCAGCCGCCCTGAGCCATCGGCTCAACCA 360
DB 401 TTGCTGTCATCGCCCTGTTGTGGGCGCAGCCGCCCTGAGCCATCGGCTCAACCA 460
QY 361 AACCCATTCGCCCTGTTGTGGGCGCAGCCGCCCTGAGCCATCGGCTCAACCA 420
DB 461 AACCCATTCGCCCTGTTGTGGGCGCAGCCGCCCTGAGCCATCGGCTCAACCA 520
QY 421 GCGGCGAGCCCGCAGGCTGTGCGGCGTGGCACTTATCGAAGCCATCGAGTCACT 480
DB 521 GCGGCGAGCCCGCAGGCTGTGCGGCGTGGCACTTATCGAAGCCATCGAGTCACT 580
QY 481 TCTCTAGTATTCGCGCGCTCTCAACCGAGTGGCAATCGAGTTCATTCCTTTGAA 540
DB 581 TCTCTAGTATTCGCGCGCTCTCAACCGAGTGGCAATCGAGTTCATTCCTTTGAA 640
QY 541 ATTAAGGCTTCGAGTTCGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 600
DB 641 ATTAAGGCTTCGAGTTCGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 700
QY 601 CGAAGCAAAAGAGATCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 660
DB 701 CGAAGCAAAAGAGATCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 760
QY 661 GTGTATTCAGGTCAGGCTTCGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 720
DB 761 GTGTATTCAGGTCAGGCTTCGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 820
QY 721 ATCCGTAATTCCT 753
DB 821 ATCCGTAATTCCT 853

RESULT 3

US-09-471-803A-6
Sequence 6, Application US/09471803A

Patent No. 6613545
GENERAL INFORMATION:
APPLICANT: KENNERNKUCHT, NICOLE
APPLICANT: SAHM, HERMANN
APPLICANT: PFEFFERLE, WALTER
TITLE OF INVENTION: NUCLEOTIDE SEQUENCE CODING FOR THE EXPORT OF
TITLE OF INVENTION: BRANCHED CHAIN AMINO ACIDS, PROCESS FOR THE
FILE REFERENCE: 21123/265496/MAS
CURRENT APPLICATION NUMBER: US/09/471, 803A
PRIOR FILING DATE: 1999-12-23
PRIOR APPLICATION NUMBER: DE 199 51 708.8
NUMBER OF SEQ ID NOS: 12
SOFTWARE: PatentIn Ver. 2.1

Tue Apr 20 06:47:16 2004

us-10-608-504-2.rn1

Page 3

SEQ ID NO 6
LENGTH: 1271
TYPE: DNA
ORGANISM: Corynebacterium glutamicum
FEATURE:
NAME/KEY: gene
LOCATION: (101)..(853)
OTHER INFORMATION: brnF
NAME/KEY: gene
LOCATION: (853)..(1176)
OTHER INFORMATION: brnE
OTHER INFORMATION: ATCC13032
US-09-471-803A-6

Query Match 100.0%; Score 753; DB 4; Length 1271;
Best Local Similarity 100.0%; Pred. No. 2,9e-231;
Matches 753; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTGCAAAAAGCGAAGATTCATTCAGCCTGAGAGGTGGCGATCCAGGCGAGCCCTG 60
DB 101 GTGCAAAAAGCGAAGATTCATTCAGCCTGAGAGGTGGCGATCCAGGCGAGCCCTG 160
QY 61 GAACGAGATGATTAAGTTATCGGCTACGAAATCGCGAAGGTCTAAAACCTCCCTT 120
DB 161 GAACGAGATGATTAAGTTATCGGCTACGAAATCGCGAAGGTCTAAAACCTCCCTT 220
QY 121 GCTGCAAGTTGGGCGATGATCCCATTTGGTATGGCTTTGGTCTCTTGGTTATTCATAC 180
DB 221 GCTGCAAGTTGGGCGATGATCCCATTTGGTATGGCTTTGGTCTCTTGGTTATTCATAC 280
QY 181 GCTGCAAGTTGGGCGAGCCCACTGTTTCCGCGCTGATTTTCGCGGCTCCACCGAA 240
DB 281 GCTGCAAGTTGGGCGAGCCCACTGTTTCCGCGCTGATTTTCGCGGCTCCACCGAA 340
QY 341 ATGCTGATGATGCGCTGCTGTTGGGCGAGCGCCCTCGGCGCATGCGCTCACCA 400
DB 401 ATGCTGATGATGCGCTGCTGTTGGGCGAGCGCCCTCGGCGCATGCGCTCACCA 460
QY 461 AACCCATTGCCCGTTCTATTGCTTTTCGCGCTTATGAGAAAGCTTACGAGTCACT 420
DB 520 AACCCATTGCCCGTTCTATTGCTTTTCGCGCTTATGAGAAAGCTTACGAGTCACT 580
QY 580 GCGGCGAGCGCGCGAGGCTGTTGGGCGAGCTTATTCATGATGAATAGCGGTTTAC 640
DB 640 GCGGCGAGCGCGCGAGGCTGTTGGGCGAGCTTATTCATGATGAATAGCGGTTTAC 700
QY 700 CGAAGCAAAAAGGAGATCCCTTCTGCTGCTGAGGTTGAGCTTACCATGCTCTT 660
DB 760 CGAAGCAAAAAGGAGATCCCTTCTGCTGCTGAGGTTGAGCTTACCATGCTCTT 720
QY 720 GTGTAATTCAGGTAAGGCGCTTATTTGGGCGCTGCTGATCTTTGGGCTCTGTGACC 780
DB 820 GTGTAATTCAGGTAAGGCGCTTATTTGGGCGCTGCTGATCTTTGGGCTCTGTGACC 840
QY 840 ATCCGAGTCTTCTTCTGGAAGGCTGCTAAA 753
DB 853 ATCCGAGTCTTCTTCTGGAAGGCTGCTAAA 853

RESULT 4
US-09-489-039A-3016

Sequence 3016, Application US/09489039A
Patent No. 6610836
GENERAL INFORMATION:
APPLICANT: Gary Breton et. al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 2709 2004001
CURRENT APPLICATION NUMBER: US/09/489,039A
CURRENT FILING DATE: 2000-01-27
PRIOR APPLICATION NUMBER: US 60/117,747
PRIOR FILING DATE: 1999-01-29
NUMBER OF SEQ ID NOS: 14342
SEQ ID NO 3016
LENGTH: 756
TYPE: DNA
ORGANISM: Klebsiella pneumoniae
US-09-489-039A-3016

Query Match 7.8%; Score 58.6; DB 4; Length 756;
Best Local Similarity 52.2%; Pred. No. 8.6e-09;
Matches 130; Conservative 0; Mismatches 119; Indels 0; Gaps 0;

QY 95 TCGCGAAGGCTTAAGAACTCCCTTCTGCTGAGGTTGGGCGATGATCCGATTTGATG 154
DB 62 TCGCGAAGGCTTAAGAACTCCCTTCTGCTGAGGTTGGGCGATGATCCGATTTGATG 121
QY 155 GCTTGGTCTCTGTTATTCATACGCTACGATGATGAGGCGAGCCCACTGTTTTCG 214
DB 122 CTTGCGATCATGATGATCCCGCTGCTTACCCCGCTGGAAGGCTGTTCTTCTCT 181
QY 215 GCTGATTTTGGGCGCTCCACGAAATGCTGCTATCGCCCTGTTGGGCGAGCCG 274
DB 182 GCATATTTACGCGCGCGCGACGACGATTTGATGATCCCGCATGTTGGCGCGAGTT 241
QY 275 CCTGGGCGCGATGCGGCTACACCATGTTGATGATGATGATGATGATGATGATGATG 334
DB 242 CGCTGAGGCGCGCGCGCTTACGATGATGATGATGATGATGATGATGATGATGATG 301
QY 335 TTTCATTCG 343
DB 302 CTTCACTGC 310

RESULT 5
US-08-232-463-14
Sequence 14, Application US/08232463
Patent No. 5670367
GENERAL INFORMATION:
APPLICANT: SCHRIFFLINER, F.
APPLICANT: FALKNER, F. G.
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,463
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935,313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6

FILING DATE: 26-AUG-1991
 ATTORNEY/AGENT INFORMATION:
 NAME: BENT, Stephen A.
 REGISTRATION NUMBER: 29,768
 REFERENCE/DOCKET NUMBER: 30472/114 IMM
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (703) 836-9300
 TELEFAX: (703) 683-4109
 TRFAX: 899149
 INFORMATION FOR SEQ ID NO: 14:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 7218 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 IMMEDIATE SOURCE:
 CLONE: PTZpdt-Fls
 US-08-232-463-14

Query Match 6.1%; Score 46.2; DB 1; Length 7218;
 Best Local Similarity 5.2%; Pred. No. 0.00027;
 Matches 21; Conservative 214; Mismatches 172; Indels 0; Gaps 0;

QY 215 GCGTATTTTCGGGCGCTCCACCGAATGCTGTCATGCGCCCTGTTGGCGGAGCGC 274
 DB 1041 GGCTGAGAGCTCGAGGAGCTTGCGATTTTTTTTTTTTTTTTTTTTTTTTTT 1100
 QY 275 CCCTGGGCGCCATCGCGTCACCAATGCTGTCGCACTTCGCGCACTTCTATGCGT 334
 DB 1101 TT 1160
 QY 335 TTTCATTCGCGCTGTCATGTCGTAATGCTGTCATGTCGTCATGTCGTCGTCGTC 394
 DB 1161 TT 1220
 QY 395 TTATCGACGAGCTGATCGCATGTCGCGCGCGCGCGCGCGCGCGCGCGCGCGC 454
 DB 1221 TT 1280
 QY 455 TTATTCATGCAATGCTGTCATGTCGTCATGTCGTCATGTCGTCATGTCGTCGTC 514
 DB 1281 TT 1340
 QY 515 CGATCGAGAGTGTGATGCTTGTGAATTAAGGCGCTGCGCTGCTGCTGCTGCT 574
 DB 1341 TT 1400
 QY 575 TTGTACGCTGACTTTGATTCCTGCGGACGAGGAGGAGGAGGAGGAGGAGGAGG 621
 DB 1401 TT 1447

RESULT 6
 US-09-107-532A-1266
 Sequence 1266, Application US/09107532A
 Patent No. 6583275
 GENERAL INFORMATION:
 APPLICANT: Lynn A Doucette-Stramm and David Bush
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
 ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
 NUMBER OF SEQUENCES: 7310
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: GENOME THERAPEUTICS CORPORATION
 STREET: 100 Beaver Street
 CITY: Waltham
 STATE: Massachusetts
 COUNTRY: USA
 ZIP: 02154
 COMPUTER READABLE FORM:
 MEDIUM TYPE: CD-ROM ISO9660
 COMPUTER: PC
 OPERATING SYSTEM: <Unknown>
 SOFTWARE: ASCII

CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/107,532A
 FILING DATE: 30-Jun-1998
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 60/085,598
 FILING DATE: 14 May 1998
 APPLICATION NUMBER: 60/051571
 FILING DATE: July 2, 1997
 ATTORNEY/AGENT INFORMATION:
 NAME: Arinello, Pamela Deneke
 REGISTRATION NUMBER: 40,489
 REFERENCE/DOCKET NUMBER: GTC-012
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (781) 893-5007
 TELEFAX: (781) 893-8277
 INFORMATION FOR SEQ ID NO: 1266:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 768 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: circular
 MOLECULE TYPE: DNA (genomic)
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 ORIGINAL SOURCE:
 ORGANISM: Enterococcus faecium
 FEATURE:
 NAME/KEY: misc_feature
 LOCATION: (B) LOCATION 1...768
 SEQUENCE DESCRIPTION: SEQ ID NO: 1266:
 US-09-107-532A-1266

Query Match 5.3%; Score 39.8; DB 4; Length 768;
 Best Local Similarity 4.83%; Pred. No. 0.0092;
 Matches 104; Conservative 0; Mismatches 107; Indels 0; Gaps 0;

QY 107 TAAACCTCCCTGCTGCAAGTTTGGCATGTAACCGATGTCATGTCGTCGTCGTC 166
 DB 92 TCAAGACAGCTACTACCTGTTTCGTTATATGCTATGACCTGCAATTTGGATCG 151
 QY 167 TGTATTCATATAGGCTGATGATGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 226
 DB 152 TTGGAGAGCTGCGGATTTATCACTATGTCGTCGTCGTCGTCGTCGTCGTCGTC 211
 QY 227 CGGCTCCACCGAATGTCGTCATGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 286
 DB 212 CTGTTTCGCGCAATTTATCACTATGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 271
 QY 287 TCGGCTACCACTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 317
 DB 272 TCGTTTCTCGACCTTCTAGTCATATCCG 302

RESULT 7
 US-09-252-991A-11376
 Sequence 11376, Application US/09252991A
 Patent No. 6551795
 GENERAL INFORMATION:
 APPLICANT: Marc J. Rubenfield et al.
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
 FILE REFERENCE: 107196.136
 CURRENT APPLICATION NUMBER: US/09/252,991A
 CURRENT FILING DATE: 1998-02-18
 PRIOR APPLICATION NUMBER: US 60/074,788
 PRIOR FILING DATE: 1998-02-18
 PRIOR APPLICATION NUMBER: US 60/094,190
 PRIOR FILING DATE: 1998-07-27
 NUMBER OF SEQ ID NOS: 33142
 SEQ ID NO: 11376
 LENGTH: 1038
 TYPE: DNA
 ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-11376

Query Match 5.2%; Score 39.2; DB 4; Length 1038;
 Best Local Similarity 55.0%; Pred. No. 0.017;
 Matches 77; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

209 TTTCCGGCCGATTTTGGCGGGCTCCACCAATGCTGTCATCGCCCTGTTGGGCG 268
 Db TGTCCGCGATGCTTTCCGCGCCGCCCACTGGTGGCATGCGATGCTCAAGGCG 532
 473 TGTCCGCGATGCTTTCCGCGCCGCCCACTGGTGGCATGCGATGCTCAAGGCG 532
 QY CAGCGCCCTGCGCGCCATCGCGCTCAACACATGCTGTGAATTCGCCACGATTTCT 328
 Db GTGCGCGGCTCTTTTCGATCATCTTCAACACCTCTCTGACCTCGAGAGACCTGCTCT 592
 QY 329 ATGCGTTTCATTTCCGCGCTG 348
 Db 593 ACGGATGAGCTTGGCGCGG 612

RESULT 8

US-09-252-991A-11418
 ; Sequence 11418, Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:

APPLICANT: Marc J. Rubenfield et al.
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 FILE REFERENCE: 107196.136
 CURRENT APPLICATION NUMBER: US/09/252,991A
 PRIOR FILING DATE: 1999-02-18
 PRIOR APPLICATION NUMBER: US 60/074,788
 PRIOR FILING DATE: 1998-02-18
 PRIOR APPLICATION NUMBER: US 60/094,190
 PRIOR FILING DATE: 1998-07-27
 NUMBER OF SEQ ID NOS: 33142
 SEQ ID NO 11418
 LENGTH: 1086
 TYPE: DNA
 ORGANISM: Pseudomonas aeruginosa
 US-09-252-991A-11418

Query Match 5.2%; Score 39.2; DB 4; Length 1086;
 Best Local Similarity 55.0%; Pred. No. 0.017; 63; Indels 0; Gaps 0;
 Matches 77; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

209 TTTCCGGCCGATTTTGGCGGGCTCCACCAATGCTGTCATCGCCCTGTTGGGCG 268
 Db 222 TGTCCGCGATGCTTTCCGCGCCGCCCACTGGTGGCATGCGATGCTCAAGGCG 281
 QY CAGCGCCCTGCGCGCCATCGCGCTCAACACATGCTGTGAATTCGCCACGATTTCT 328
 Db 282 GTGCGCGGCTCTTTTCGATCATCTTCAACACCTCTCTGACCTCGAGAGACCTGCTCT 341
 QY 329 ATGCGTTTCATTTCCGCGCTG 348
 Db 342 ACGGATGAGCTTGGCGCGG 361

RESULT 9

US-09-621-976-8976/c
 ; Sequence 8976, Application US/09621976
 ; Patent No. 6639063
 ; GENERAL INFORMATION:

APPLICANT: Dumas Milne Edwards, J.B.
 APPLICANT: Joubert, S.
 TITLE OF INVENTION: ESTE and Encoded Human Proteins.
 FILE REFERENCE: GENSET.054PR2
 CURRENT APPLICATION NUMBER: US/09/621,976
 CURRENT FILING DATE: 2000-07-21
 NUMBER OF SEQ ID NOS: 19335
 SOFTWARE: Patent.pm
 SEQ ID NO 8976

LENGTH: 399
 TYPE: DNA
 ORGANISM: Homo sapiens
 US-09-621-976-8976

Query Match 5.2%; Score 39; DB 4; Length 399;
 Best Local Similarity 12.6%; Pred. No. 0.012;
 Matches 30; Conservative 112; Mismatches 97; Indels 0; Gaps 0;

QY 509 GAGTGGCGATTCGAGAGTTCCTTTGAAATTAAGGCGCTCGAGTTCGCGCTTGGT 568
 Db 242 SAAMRRKKMTCKMGRSWSRSTGYAMWKSWCTSRKMTYKRRKKRKTSTRT 183
 QY CTCTCTTTGTCACGCTGATTTGATTCCTGCGCAAGAAAAGAGATCCTTCTTGC 628
 Db 182 CYRGSYKCMKAYYTKRRKRTWTYYKSYMSKMTWRMTAYYWTBRMTRTKMTW 123
 QY 629 TGCTCGAGGTTGAGCTTCACCATTCCTTGTGTATTCAGGTCAGGCGCTATTG 688
 Db 122 CTMCMKCTTYMAGTMYRRTYRYAKAKMSKCTWSITCYCMKYMAKCMSTWMS 63
 QY 689 CGGCGCTGCTGATCTTCTGGCTGTGACCATCCGACTTCTTCTGGAAGGCT 747
 Db 62 MMKMGKSMWKMTYYYYMMKMSKMTYMSVCYARKCMRTYAKTYMTTCMTWKR 4

RESULT 10

US-09-252-991A-3205/c
 ; Sequence 3205, Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:

APPLICANT: Marc J. Rubenfield et al.
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 FILE REFERENCE: 107196.136
 CURRENT APPLICATION NUMBER: US/09/252,991A
 PRIOR FILING DATE: 1999-02-18
 PRIOR APPLICATION NUMBER: US 60/074,788
 PRIOR FILING DATE: 1998-02-18
 PRIOR APPLICATION NUMBER: US 60/094,190
 PRIOR FILING DATE: 1998-07-27
 NUMBER OF SEQ ID NOS: 33142
 SEQ ID NO 3205
 LENGTH: 1140
 TYPE: DNA
 ORGANISM: Pseudomonas aeruginosa
 US-09-252-991A-3205

Query Match 4.9%; Score 36.8; DB 4; Length 1140;
 Best Local Similarity 58.0%; Pred. No. 0.1;
 Matches 65; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

QY 211 TCCGCGCTGATTTTGGCGGGCTCCACCAATGCTGTCATCGCCCTGTTGGGCGCA 270
 Db 465 TACTTCCTGATCAACCTCGCGCGCCCTCGCGCTGATCGAGCTTCGGGCTC 406
 QY 271 GCGCCCTGCGCGCCATCGCGCTCAACCATGCTGTGTAAGTTCGCCACG 322
 Db 405 GCGGCCAGGCGGCACTTCTCTGTGTCACCGCGCTGTCTACTTCTCTACG 354

RESULT 11

US-09-252-991A-2761
 ; Sequence 2761, Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:

APPLICANT: Marc J. Rubenfield et al.
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 FILE REFERENCE: 107196.136
 CURRENT APPLICATION NUMBER: US/09/252,991A
 CURRENT FILING DATE: 1999-02-18
 PRIOR APPLICATION NUMBER: US 60/074,788

PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO: 2761
LENGTH: 2061
TYPE: DNA
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-2761

Query Match 4.9%; Score 36.8; DB 4; Length 2061;
Best Local Similarity 58.0%; Pred. No. 0.14; Indels 0; Gaps 0;
Matches 65; Conservative 0; Mismatches 47;

QY 211 TCGGCGCTGATTTTCGGCGGCTCCACCAATGCTGTCATCGCCCTGTTGTGGCGCA 270
DB 1273 TACTTCTGATCAACTCGGCGCGCCCTCGGCGCGCTGATCGGCTGACTCTCGCGCTC 1332
QY 271 GCGCCCTCGGCGCGCTCCGCGCTCACCAATGCTGTCGTAACCTTCGCGCAGC 322
DB 1333 GCGCGCCAGCGCGCACTTCTCGTCAACCGCGCTGCTACTTCTGCTACG 1384

RESULT 12
US-09-020-956-12/C
Sequence 12, Application US/09020956
Patent No. 6261562
GENERAL INFORMATION:
APPLICANT: Xu, Jiangchun
APPLICANT: Dillon, David C.
TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY OF PROSTATE CANCER AND METHODS FO
NUMBER OF SEQUENCES: 178
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/020,956
FILING DATE: 09-FEB-1998
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Makl, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 210121.427C2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 751 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-09-020-956-12

Query Match 4.8%; Score 36.4; DB 3; Length 751;
Best Local Similarity 59.8%; Pred. No. 0.11; Indels 39; Gaps 0;
Matches 58; Conservative 0; Mismatches 39;

QY 227 CGGGCTCCACCAATGCTGTCATCGCCCTGTTGTGGCGCAGCGCCCTTGCGCGCA 286
DB 410 CGTCTTCTTCACTGCTGTCATGCTGAGGNGCAGCTGCTGTGCTGCT 351
QY 287 TCGGCGTCAACCAATGCTGTCGTAACCTTCGCGCAGT 323

DB 350 TGGTGTACCAACATGCTGAGCACTTCCCTGACCT 314

RESULT 13
US-09-030-607-12/C
Sequence 12, Application US/09030607
Patent No. 6262245
GENERAL INFORMATION:
APPLICANT: Xu, Jiangchun
APPLICANT: Dillon, David C.
TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY OF PROSTATE CANCER AND METHODS FO
NUMBER OF SEQUENCES: 224
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/030,607
FILING DATE: 25-FEB-1998
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Makl, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 210121.427C3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 751 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-09-030-607-12

Query Match 4.8%; Score 36.4; DB 3; Length 751;
Best Local Similarity 59.8%; Pred. No. 0.11; Indels 39; Gaps 0;
Matches 58; Conservative 0; Mismatches 39;

QY 227 CGGGCTCCACCAATGCTGTCATCGCCCTGTTGTGGCGCAGCGCCCTTGCGCGCA 286
DB 410 CGTCTTCTTCACTGCTGTCATGCTGAGGNGCAGCTGCTGTGCTGCT 351
QY 287 TCGGCGTCAACCAATGCTGTCGTAACCTTCGCGCAGT 323
DB 350 TGGTGTACCAACATGCTGAGCACTTCCCTGACCT 314

RESULT 14
US-09-439-313-12/C
Sequence 12, Application US/09439313
Patent No. 6329505
GENERAL INFORMATION:
APPLICANT: Xu, Jiangchun
APPLICANT: Dillon, David C.
APPLICANT: Mitcham, Jennifer L.
APPLICANT: Harlocke, Susan Louise
APPLICANT: Jiang Yugu
APPLICANT: Reed, Steven G.
APPLICANT: Kalos, Michael
APPLICANT: Fanger, Gary
APPLICANT: Retter, Mark
APPLICANT: Solk, John

```
APPLICANT: Day, Craig
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
FILE REFERENCE: 210121.427C9
CURRENT APPLICATION NUMBER: US/09/439,313
NUMBER OF SEQ ID NOS: 575
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 12
LENGTH: 751
TYPE: DNA
ORGANISM: Homo sapien
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)..(751)
OTHER INFORMATION: n = A,T,C or G
US-09-439-313-12
```

```
Query Match          4.8%; Score 36.4; DB 4; Length 751;
Best Local Similarity 59.8%; Pred. No. 0.11;
Matches 58; Conservative 0; Mismatches 39; Indels 0; Gaps 0;
```

```
QY 227 CGGGCTCCACCGAAATGCTGTCATCGCCCTGTTGTGGGCGACGCGCCCTGGGCGCCA 286
DB 410 CGTTCTTCTTCATCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCT 351
QY 287 TCGCGCTACCCACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 323
DB 350 TGGTGTACACCAATGGCTGAGCACTTCCTGACGT 314
```

RESULT 15

```
US-09-352-616A-12/C
Sequence 12, Application US/09352616A
Patent No. 6395278
GENERAL INFORMATION:
APPLICANT: Dillon, Davin C.
APPLICANT: Harlocker, Susan Louise
APPLICANT: Jiang, Yugu
APPLICANT: Xu, Jiangchun
APPLICANT: Mitcham, Jennifer Lynn
TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS
FILE REFERENCE: 210121.427C8
CURRENT APPLICATION NUMBER: US/09/352,616A
NUMBER OF SEQ ID NOS: 472
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 12
LENGTH: 751
TYPE: DNA
ORGANISM: Homo sapien
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)..(751)
OTHER INFORMATION: n = A,T,C or G
US-09-352-616A-12
```

```
Query Match          4.8%; Score 36.4; DB 4; Length 751;
Best Local Similarity 59.8%; Pred. No. 0.11;
Matches 58; Conservative 0; Mismatches 39; Indels 0; Gaps 0;
```

```
QY 227 CGGGCTCCACCGAAATGCTGTCATCGCCCTGTTGTGGGCGACGCGCCCTGGGCGCCA 286
DB 410 CGTTCTTCTTCATCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCT 351
QY 287 TCGCGCTACCCACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 323
DB 350 TGGTGTACACCAATGGCTGAGCACTTCCTGACGT 314
```


NUMBER OF SEQUENCES: 7310
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: GENOME THERAPEUTICS CORPORATION
 STREET: 100 Beaver Street
 CITY: Waltham
 STATE: Massachusetts
 COUNTRY: USA
 ZIP: 02354
 COMPUTER READABLE FORM:
 MEDIUM TYPE: CD-ROM ISO9660
 COMPUTER: PC
 OPERATING SYSTEM: <Unknown>
 SOFTWARE: ASCII
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/10/417,884A
 FILING DATE: 17-Apr-2003
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/09/107,532
 FILING DATE: 30-Jun-1998
 APPLICATION NUMBER: 60/085,598
 FILING DATE: 14 May 1998
 APPLICATION NUMBER: 60/051571
 FILING DATE: July 2, 1997
 ATTORNEY/AGENT INFORMATION:
 NAME: Ariniello, Pamela Deneka
 REGISTRATION NUMBER: 40,489
 REFERENCE/DOCKET NUMBER: GTC-012
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (781)893-5007
 TELEFAX: (781)893-8277
 INFORMATION FOR SEQ ID NO: 1266:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 768 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: circular
 MOLECULE TYPE: DNA (genomic)
 HYPOTHEetical: NO
 ANTI-SENSE: NO
 ORIGINAL SOURCE:
 ORGANISM: Enterococcus faecium
 FEATURE:
 NAME/KEY: mlec_feature
 LOCATION: (B) LOCATION 1...768
 SEQUENCE DESCRIPTION: SEQ ID NO: 1266:
 US-10-417-884A-1266

Query Match 5.3%; Score 39.8; DB 6; Length 768;
 Best Local Similarity 49.3%; Pred. No. 0.0082;
 Matches 104; Conservative 0; Mismatches 107; Indels 0; Gaps 0;
 QY 107 TAAAAACCTCCCTGCTGAGGTTGGGATGATACCGGATTTGGTATTCGTTGGTCTCT 166
 DB 92 TCAAGACACGTTACTTACCTGTTTCGTTATGATGATGACATTGGATCGATTGGATCG 151
 QY 167 TGGTTATTCANAGGCGTACGATGAGATGGTGGGACGCCCACTGTTTCCGGCTGATTTCC 226
 DB 152 TTGGAAAGCTGCCGATTTCAATCATCATGATGCTGACGTTATCTCCCTGGTGTATG 211
 QY 227 CGGGCTCCACGGAATGCTGTCATCGCCCTGTTGGGCGCAGCGCCCTGGGCGCCA 286
 DB 212 CTGGTTTCCCAATTATCATCAGTCAGCATGCTTGCTAGTCACAGCCCATTTGTTTCCA 271
 QY 287 TCGGCTCAGCAGATTGCTGATGATGATCCG 317
 DB 272 TCGTTTCTCGACCTTTCTAGCATTCCTCG 302

RESULT 3
 US-10-767-701-5263
 Sequence 5263, Application US/10767701
 GENERAL INFORMATION:
 APPLICANT: Kovalic, David K.

APPLICANT: Zhou, Yihua
 APPLICANT: Cao, Yongwei
 TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
 TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
 FILE REFERENCE: 38-21(5353)B
 CURRENT APPLICATION NUMBER: US/10/767,701
 CURRENT FILING DATE: 2004-01-29
 NUMBER OF SEQ ID NOS: 63128
 SEQ ID NO 5263
 LENGTH: 619
 TYPE: DNA
 ORGANISM: Sorghum bicolor
 FEATURE:
 OTHER INFORMATION: Clone ID: SORBI-28MAY03-CLUS11974_1
 US-10-767-701-5263

Query Match 4.7%; Score 35.4; DB 6; Length 619;
 Best Local Similarity 54.1%; Pred. No. 0.19; 61; Indels 0; Gaps 0;
 Matches 72; Conservative 0; Mismatches 61; Indels 0; Gaps 0;
 QY 226 GCGGCTCCACCGAATGCTGTCATGCGCCCTGTTGTGGGCGCAGCGCCCTGGGCGCC 285
 DB 132 GCGGCTTCACCTTGTGCTGTGCTGCGGGCTGGCGGTGACGGGACAGTGTGGGCTC 191
 QY 286 ATCGGCTCCACCATTTGCTGATGATGATGATGATGATGATGATGATGATGATGATG 345
 DB 192 ATCGGCGACGCGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATG 251
 QY 346 CTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 358
 DB 252 GTGGCGCTGCTCA 264

RESULT 4
 US-10-767-701-10866
 Sequence 10866, Application US/10767701
 GENERAL INFORMATION:
 APPLICANT: Kovalic, David K.
 APPLICANT: Zhou, Yihua
 APPLICANT: Cao, Yongwei
 TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
 TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
 FILE REFERENCE: 38-21(5353)B
 CURRENT APPLICATION NUMBER: US/10/767,701
 CURRENT FILING DATE: 2004-01-29
 NUMBER OF SEQ ID NOS: 63128
 SEQ ID NO 10866
 LENGTH: 1034
 TYPE: DNA
 ORGANISM: Sorghum bicolor
 FEATURE:
 OTHER INFORMATION: Clone ID: SORBI-28MAY03-CLUS4850_1
 US-10-767-701-10866

Query Match 4.6%; Score 34.6; DB 6; Length 1034;
 Best Local Similarity 49.7%; Pred. No. 0.45; 89; Indels 0; Gaps 0;
 Matches 88; Conservative 0; Mismatches 89; Indels 0; Gaps 0;
 QY 241 ATGCTGTCATGACCTGCTGTTGTGGGCGCAGCGCCCTGGGCGCATGCGCTCACACA 300
 DB 151 ATGCGGCGCCCTTACTTACGCGGAGGCGCCCTGGGCGCACCGGCGCGCGCGCCCT 210
 QY 301 TTGCTGTAACTTCCGCGCAGTATTTATGCTTTTATTTCCGCTGATGATGATGATG 360
 DB 211 GTGGCGGTGATGCTGCGCGAGTTTGGCGCGGCTGATGATGATGATGATGATGATG 270
 QY 361 AACCCATGCGCGCTTATGCTGATGATGATGATGATGATGATGATGATGATGATG 417
 DB 271 AAGGCATGACCTCTCCGACGCGGACCTTCAACGATCAGCAGGCGGCGCGCGCTC 327

RESULT 5
 US-10-100-683-983/c

```
; Sequence 983, Application US/10100683
; GENERAL INFORMATION:
; APPLICANT: Rosen, et al.
; TITLE OF INVENTION: Human Secreted Proteins
; FILE REFERENCE: PS900
; CURRENT APPLICATION NUMBER: US/10/100,683
; PRIOR FILING DATE: 2002-03-19
; PRIOR APPLICATION NUMBER: US 60/040,162
; PRIOR FILING DATE: 1997-03-07
; PRIOR APPLICATION NUMBER: US 60/043,576
; PRIOR FILING DATE: 1997-04-11
; PRIOR APPLICATION NUMBER: US 60/047,601
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: US 60/056,845
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: US 60/043,314
; PRIOR FILING DATE: 1997-04-11
; PRIOR APPLICATION NUMBER: US 60/047,599
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: US 60/056,664
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: US 60/043,314
; PRIOR FILING DATE: 1997-04-11
; PRIOR APPLICATION NUMBER: US 60/047,632
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: US 60/056,892
; PRIOR FILING DATE: 1997-08-22
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 13468
; SOFTWARE: Patent Ver. 2.0
; SEQ ID NO 983
; LENGTH: 3173
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-100-683-983

Query Match          4.6%; Score 34.6; DB 6; Length 3173;
Best Local Similarity 75.4%; Pred. No. 0.79;
Matches 43; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 244 CCGGTCATCGCCCTGCTGTGGGCGGCGCCCTGGGGCGCATGCGGTCAACCA 300
DB 2360 CAGCTCCGCGCCCGCAGTGTGGGCTCAGGGCTCTGGGCGCCATGGGTCAACCA 2304

RESULT 6
US-10-100-683-11511
; Sequence 11511, Application US/10100683
; GENERAL INFORMATION:
; APPLICANT: Rosen, et al.
; TITLE OF INVENTION: Human Secreted Proteins
; FILE REFERENCE: PS900
; CURRENT APPLICATION NUMBER: US/10/100,683
; PRIOR FILING DATE: 2002-03-19
; PRIOR APPLICATION NUMBER: US 60/040,162
; PRIOR FILING DATE: 1997-03-07
; PRIOR APPLICATION NUMBER: US 60/043,576
; PRIOR FILING DATE: 1997-04-11
; PRIOR APPLICATION NUMBER: US 60/047,601
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: US 60/056,845
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: US 60/043,580
; PRIOR FILING DATE: 1997-04-11
; PRIOR APPLICATION NUMBER: US 60/047,599
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: US 60/056,664
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: US 60/043,314
; PRIOR FILING DATE: 1997-04-11
; PRIOR APPLICATION NUMBER: US 60/047,632
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: US 60/056,892
```

```
; PRIOR FILING DATE: 1997-08-22
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 13468
; SOFTWARE: Patent Ver. 2.0
; SEQ ID NO 11511
; LENGTH: 7071
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-100-683-11511

Query Match          4.6%; Score 34.6; DB 6; Length 7071;
Best Local Similarity 75.4%; Pred. No. 1.2;
Matches 43; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 244 CCGGTCATCGCCCTGCTGTGGGCGGCGCCCTGGGGCGCATGCGGTCAACCA 300
DB 785 CAGCTCCGCGCCCGCAGTGTGGGCTCAGGGCTCTGGGCGCCATGGGTCAACCA 841

RESULT 7
US-10-767-701-7545
; Sequence 7545, Application US/10767701
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(5353)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; PRIOR FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 7545
; LENGTH: 666
; TYPE: DNA
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-CLUS92927_1
; US-10-767-701-7545

Query Match          4.6%; Score 34.4; DB 6; Length 666;
Best Local Similarity 57.4%; Pred. No. 0.42;
Matches 62; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

QY 215 GCGTGAATTTGCGGGGCTCCACCGAAATGCTGTGATCGCCCTGCTGTGGGCGGCGGCGG 274
DB 31 GCCCGACCCACACCGGCGGAGAGAGACCCGCTCATCAACCCGTTGCGGAGCGGCGCGCGC 90

QY 275 CCCGCGGGCGCGCATGCGGCTGACCAACATTGCTGGTGAACCTTCGCGCAGC 322
DB 91 CGCTGGCGTCCCTGGCGTGAACCGCGCTGCTGTGTGCGGTGCGGCGAG 138

RESULT 8
US-10-767-701-744/C
; Sequence 744, Application US/10767701
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(5353)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; PRIOR FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 744
; LENGTH: 561
; TYPE: DNA
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-CLUS124346_1
; US-10-767-701-744
```

Query Match 4.5%; Score 34; DB 6; Length 561;
Best Local Similarity 54.9%; Pred. No. 0.53;
Matches 67; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

QY 202 CCACTGTTTCCGCGCTGATTTTCGCGGCTCCACCGAATGCTGTCATCGCCCTGTT 261
DB 333 CCAAGGAGCGGAGACCCCTGAGCTTCTTGAGCCGAGCCGATCCTGGGCTGCTCTGTC 274
QY 262 GTGGGCGGAGCGCCCTTGGGCGCCATCGGCTCACACATGCTGTGAATTCGGCCAC 321
DB 273 ACCAGCCCGGAGCGGCTGCGGCTCTCTCTCCAGACCTTGATGATGAGTACTCGCC 214
QY 322 GT 323
DB 213 CT 212

RESULT 9

US-10-765-790-62
Sequence 62, Application US/10765790
GENERAL INFORMATION:
APPLICANT: Bayer Healthcare LLC
APPLICANT: Beard, Chris
APPLICANT: Burgess, Chris
APPLICANT: Gannon, Allison
APPLICANT: Harvey, Jeanne
APPLICANT: Lechner, John F.
APPLICANT: Li, Zheng
TITLE OF INVENTION: Identification and Verification of Methylation Marker Sequences
FILE REFERENCE: 1657/2035
CURRENT APPLICATION NUMBER: US/10/765,790
CURRENT FILING DATE: 2004-01-27
PRIOR APPLICATION NUMBER: US 10/737,082
PRIOR FILING DATE: 2003-12-16
NUMBER OF SEQ ID NOS: 300
SOFTWARE: PatentIn version 3.2
SEQ ID NO 62
LENGTH: 23490
TYPE: DNA
ORGANISM: Homo sapiens
US-10-765-790-62

Query Match 4.5%; Score 33.8; DB 6; Length 23490;
Best Local Similarity 49.7%; Pred. No. 3.8;
Matches 86; Conservative 0; Mismatches 87; Indels 0; Gaps 0;

QY 201 CCCACTGTTTCCGCGCTGATTTTCGCGGCTCCACCGAATGCTGTCATCGCCCTGTT 260
DB 21210 CTGGTGTCTCTCCCTCCCTGCGATCTGTGTCACCTTGTGCTGTGAGGCTGGGA 21269
QY 261 TGTGGGCGGAGCGCCCTTGGGCGCCATCGGCTCACACATGCTGTGAATTCGGCCA 320
DB 21270 GGAAGCTGGGCGGCTGCGGCTGCTGAGCGGCTATGCGCTGTGGGCGCTTGACT 21329
QY 321 CGATTTGATGCGTTTCATTTCCGCTGATGATGATGATGATGATGATGATGATGATG 373
DB 21330 CCTCTACTGCT 21382

RESULT 10
US-60-548-091-5686
Sequence 5686, Application US/60548091
GENERAL INFORMATION:
APPLICANT: CARGILL, Michele et al.
TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
TITLE OF INVENTION: STROKE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CU001506
CURRENT APPLICATION NUMBER: US/60/548,091
CURRENT FILING DATE: 2004-02-27
NUMBER OF SEQ ID NOS: 24433
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 5686

LENGTH: 24477
TYPE: DNA
ORGANISM: Homo sapiens
US-60-548-091-5686

Query Match 4.5%; Score 33.8; DB 7; Length 24477;
Best Local Similarity 51.9%; Pred. No. 3.9;
Matches 68; Conservative 3; Mismatches 60; Indels 0; Gaps 0;

QY 242 TGTGTGATGAGCTGCTGCTGTTGTGGGCGGAGGCGCCCTTGGGCGCATGCGCTACCAAT 301
DB 17603 TGTGTGATGAGCTGCTGCTGTTGTGGGCGGAGGCGCCCTTGGGCGCATGCGCTACCAAT 17662
QY 302 TGTGTGATGAGCTGCTGCTGTTGTGGGCGGAGGCGCCCTTGGGCGCATGCGCTACCAAT 361
DB 17663 GCTTGAAGAGACACCGGAGCATTTCAATCAGTGGGTTCCGCTGCTGATTTACC 17722
QY 362 ACCCATTTGCC 372
DB 17723 TGCCCTTGCC 17733

RESULT 11

US-09-804-291A-500
Sequence 500, Application US/09804291A
GENERAL INFORMATION:
APPLICANT: ZOZULA, SERGEY
TITLE OF INVENTION: HUMAN OLFACTORY RECEPTORS AND GENES ENCODING SAME
FILE REFERENCE: 100337,54287US
CURRENT APPLICATION NUMBER: US/09/804,291A
CURRENT FILING DATE: 2001-03-13
PRIOR APPLICATION NUMBER: 60/188,914
PRIOR FILING DATE: 2000-03-13
PRIOR APPLICATION NUMBER: 60/192,033
PRIOR FILING DATE: 2000-03-24
PRIOR APPLICATION NUMBER: 60/198,474
PRIOR FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/199,335
PRIOR FILING DATE: 2000-04-24
PRIOR APPLICATION NUMBER: 60/207,702
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/213,849
PRIOR FILING DATE: 2000-06-23
PRIOR APPLICATION NUMBER: 60/226,534
PRIOR FILING DATE: 2000-08-16
PRIOR APPLICATION NUMBER: 60/230,732
PRIOR FILING DATE: 2000-09-07
PRIOR APPLICATION NUMBER: 60/266,862
PRIOR FILING DATE: 2001-02-07
NUMBER OF SEQ ID NOS: 531
SOFTWARE: PatentIn Ver. 3.2
SEQ ID NO 500
LENGTH: 930
TYPE: DNA
ORGANISM: Homo sapiens
US-09-804-291A-500

Query Match 4.4%; Score 33.2; DB 5; Length 930;
Best Local Similarity 54.0%; Pred. No. 1.2;
Matches 68; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

QY 467 AATAGCGTTTCTACTCTGAGGATTTGCGGCTTTCACCGAGTGGCGATCGAGAGT 526
DB 509 AATATGATACATACATTAATGATGATGATGATGATGATGATGATGATGATGATGATG 568
QY 527 TGATTCCTTTGAATTAAGGCGCTGAGTTCGCGCTTGTCTCTCTTTGTCAAGCTGA 586
DB 569 CAAGCTCTTGAATTAAGTATGATGATGATGATGATGATGATGATGATGATGATGATG 628
QY 587 CTTTGG 592
DB 629 TGCTGG 634

RESULT 12
US-60-548-091-19934
; Sequence 19934, Application US/60548091
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: CL001506
; CURRENT APPLICATION NUMBER: US/60/548,091
; CURRENT FILING DATE: 2004-02-27
; NUMBER OF SEQ ID NOS: 24433
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19934
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-60-548-091-19934

Query Match 4.4%; Score 33; DB 7; Length 201;
Best Local Similarity 52.7%; Pred. No. 0.67;
Matches 69; Conservative 1; Mismatches 61; Indels 0; Gaps 0;
DB 242 TGCTGATCATGCGCCCTGTTGTGGGCGCAGCGCCCTGGGCGGCATGCGCTCACCACAT 301
46 TGCTGATCATGCGCTCCCTGTTGTGGGCGCAGCGCCCTGGGCGGCATGCGCTCACCACAT 105
QY 302 TGCTGATCATGCGCCCTGTTGTGGGCGCAGCGCCCTGGGCGGCATGCGCTCACCACAT 361
106 GCTTCAGAGACACCGCGCACATTTCAATCCATGCGGTTCCGCTGCTGGGATTCCTAC 165
DB 362 ACCCATTTGCC 372
QY 166 TGCCCTTGCC 176

RESULT 13
US-60-548-091-19938
; Sequence 19938, Application US/60548091
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: CL001506
; CURRENT APPLICATION NUMBER: US/60/548,091
; CURRENT FILING DATE: 2004-02-27
; NUMBER OF SEQ ID NOS: 24433
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19938
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-60-548-091-19938

Query Match 4.4%; Score 33; DB 7; Length 201;
Best Local Similarity 52.7%; Pred. No. 0.67;
Matches 69; Conservative 1; Mismatches 61; Indels 0; Gaps 0;
DB 242 TGCTGATCATGCGCCCTGTTGTGGGCGCAGCGCCCTGGGCGGCATGCGCTCACCACAT 301
49 TGCTGATCATGCGCTCCCTGTTGTGGGCGCAGCGCCCTGGGCGGCATGCGCTCACCACAT 108
QY 302 TGCTGATCATGCGCCCTGTTGTGGGCGCAGCGCCCTGGGCGGCATGCGCTCACCACAT 361
109 GCTTCAGAGACACCGCGCACATTTCAATCCATGCGGTTCCGCTGCTGGGATTCCTAC 168
DB 362 ACCCATTTGCC 372
QY 169 TGCCCTTGCC 179

RESULT 14
US-10-767-701-9054

; Sequence 9054, Application US/10767701
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(5353)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 9054
; LENGTH: 1028
; TYPE: DNA
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-CLUS73927_1
US-10-767-701-9054

Query Match 4.4%; Score 32.8; DB 6; Length 1028;
Best Local Similarity 54.0%; Pred. No. 1.7;
Matches 67; Conservative 0; Mismatches 57; Indels 0; Gaps 0;
DB 609 AAGCAGATCCCTTCTCTGCTGCTGCGCAGGTTGAGCTTCACATGCTCTTGTGTAAT 668
746 ATAGCCGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 805
QY 669 TCCAGGTAGGCGCTATTGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 728
806 GCCCAGATGAGCGCACGCGGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 865
DB 729 CTTG 732
QY 866 GTCC 869

RESULT 15
US-10-767-701-23189
; Sequence 23189, Application US/10767701
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(5353)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 23189
; LENGTH: 519
; TYPE: DNA
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: 14593439
US-10-767-701-23189

Query Match 4.3%; Score 32.6; DB 6; Length 519;
Best Local Similarity 52.6%; Pred. No. 1.4;
Matches 71; Conservative 0; Mismatches 64; Indels 0; Gaps 0;
DB 182 GCTAGATGATGAGGCGCAGCCCATGTTTCCGCTGATTTTGGCGGCTCCACCGAAA 241
70 GCTAGGCGGCGGAGACCGCATGCTGCTTCCGCGCGCGGCTCTTCCGCG 129
QY 242 TGCTGATCATGCGCCCTGTTGTGGGCGCAGCGCCCTGGGCGGCATGCGCTCACCACAT 301
130 CCTGCTGCGCGCGCTGCGCAGAGCGCGCGCGCTTGGCCCGCGCGCGCGCGCGCGCGCG 189
DB 302 TGCTGATCATGCT 316
QY 190 CGTGCATCAAGTAC 204

Tue Apr 20 06:47:17 2004

us-10-608-504-2.rnpn

Page 6

Search completed: April 17, 2004, 20:12:58
Job time : 88.0151 secs

Tue Apr 20 06:47:17 2004

us-10-608-504-2.rst

Page 1

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 17, 2004, 06:48:18 ; Search time 2438.38 Seconds
(without alignments)
9221.792 Million cells updates/sec

Title: US-10-608-504-2
Perfect score: 753
Sequence: 1 gtgcacaaacgcagagat.....tcttggaagagctgctctaaa 753

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 27513289 segs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estinu:*
4: em_estmu:*
5: em_estov:*
6: em_estp1:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_ges_hum:*
18: em_ges_inv:*
19: em_ges_pin:*
20: em_ges_vrt:*
21: em_ges_fun:*
22: em_ges_mam:*
23: em_ges_mus:*
24: em_ges_pro:*
25: em_ges_rod:*
26: em_ges_png:*
27: em_ges_vr1:*
28: gb_ges1:*
29: gb_ges2:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	54.6	7.3	473	28	AQ936633 HSJ41-905
2	45	6.0	1201	13	BX356664 BX356664
3	39.4	743	29	CC684922	CC684922 OGUAX55TH
4	39.4	787	29	CNS009WA	AL054013 Drosoph11

5	39.2	5.2	1201	13	BX381961	BX381961
6	39	5.2	925	23	CNS0091P	AL053013 Drosoph11
7	38.6	5.1	1201	13	BX361080	BX361080 BX361080
8	38	5.0	835	14	CB657749	CB657749 OSUNEC13F
9	37.4	5.0	697	28	BH887860	BH887860 LB01712A.
10	37.4	5.0	1108	13	BX382084	BX382084 BX382084
11	37.2	4.9	1308	12	BG845919	BG845919 1024011FO
12	37.2	4.9	1454	28	BR101158	BR101158 60175473
13	36.8	4.9	1464	28	BH805539	BH805539 1008061A1
14	36.8	4.9	1201	29	CNS016BR	AL106545 Drosoph11
15	36.6	4.9	539	10	BE013856	BE013856 125354 MA
16	36.6	4.9	942	29	CNS00CT2	AL059804 Drosoph11
17	36.4	4.8	298	12	BJ525053	BJ525053 BJ525053
18	36.2	4.8	673	12	BJ287010	BJ287010 BJ287010
19	36.2	4.8	842	13	BX311708	BX311708 BX311708
20	36.2	4.8	902	29	CG358991	CG358991 OGAU35TH
21	36.2	4.8	1101	13	BQ678281	BQ678281 AGRICOURT
22	36	4.8	559	13	BX384405	BX384405 BX384405
23	36	4.8	666	28	CC170529	CC170529 1197407.b
24	36	4.8	784	29	CG250542	CG250542 OXEU07TV
25	36	4.8	836	29	CG449984	CG449984 OGVH07TV
26	36	4.8	862	28	BZ644270	BZ644270 OGCBX45TC
27	36	4.8	863	29	CG449978	CG449978 OGVH07TV
28	35.8	4.8	410	9	A1138160	A1138160 TUBS23 T
29	35.8	4.8	641	28	BZ375719	BZ375719 1e60910.9
30	35.8	4.8	641	29	CC668174	CC668174 OGUY196TV
31	35.8	4.8	696	29	CC654373	CC654373 OGVH12TV
32	35.8	4.8	980	29	CG326716	CG326716 OG2C063TV
33	35.6	4.7	327	10	BF543765	BF543765 UI-R-C3-t
34	35.6	4.7	325	13	BX424877	BX424877 BX424977
35	35.6	4.7	532	9	A1293892	A1293892 LP07146.5
36	35.6	4.7	657	13	CA109167	CA109167 SCGHR106
37	35.6	4.7	683	14	CF850119	CF850119 pEMu013XA
38	35.6	4.7	694	28	CC337836	CC337836 OGTA14TV
39	35.6	4.7	798	12	BG855128	BG855128 1024041HO
40	35.6	4.7	900	29	CG371461	CG371461 OG2A005TV
41	35.6	4.7	977	29	CNS00UX7	AL076850 Drosoph11
42	35.6	4.7	985	13	BX359916	BX359916 BX359916
43	35.6	4.7	1201	13	BX381971	BX381971 BX381971
44	35.4	4.7	398	14	CF865113	CF865113 pS25014XB
45	35.4	4.7	454	9	A1575130	A1575130 UI-R-G0-u

ALIGNMENTS

RESULT 1	AQ936633	473 bp	DNA	linear	GSS 23-AUG-2000
LOCUS	HSJ41-905	Human	NotI clones	Homo sapiens genomic	genomic survey
DEFINITION	HSJ41-905 Human	NotI clones	Homo sapiens	genomic	genomic survey
ACCESSION	AQ936633				
VERSION	AQ936633.1	GI:7213011			
KEYWORDS	GSS.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
AUTHORS	Zabarovskiy,E.R., Gazatullin,R., Podowski,R.M., Zabarovska,V.V., Xie,L., Muravenko,O.V., Kozlyev,S., Petrenko,L., Skobeleva,N., Li,J., Protodopov,A., Kashuba,V., Ernberg,I., Winberg,G. and Wahlstedt,C.				
TITLE	NotI clones in the analysis of the human genome				
JOURNAL	Nucleic Acids Res.	28 (7)			1635-1639 (2000)
MEDLINE	20175728				
PUBMED	10710430				
COMMENT	Contact: Podowski RM Center for Genomics Research Karolinska Institute 17177 Stockholm, Sweden Tel: +46-8-728-6372 Fax: +46-8-337983				

Email: Raf.Podowski@agr.ki.se
Class: NotI site.
Location/Qualifiers

FEATURES
source
1..473
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone_id="Human NotI clones"

ORIGIN

Query Match
Best Local Similarity 56.2%; Pred. No. 0.00068;
Matches 99; Conservative 0; Mismatches 77; Indels 0; Gaps 0;

QY 193 TGGGAGCCCGCCACTGTTTCCGCGCTGATTTCCGGGGCTCCACCGAAATGCTGTCATC 252
DB 283 TGGGAGGCGCCAGGCGCTCTCGGCATCTCTNCGGGGGCGCGCCAGTTGGTGGCCATC 342
QY 253 GCGCTGCTGTGGGGCGGAGGCGCCCTGGGCGCATCGCGTCACACATGCTGTGAAC 312
DB 343 GGCATCTTAAAGCGCGAGCAACTGCTGTGATCTCTACACACACTGCTGTACT 402
QY 313 TTCCGCGCAGTATTCATGCTTTTCATTCGCGCTGATGTGTAACCCCAT 368
DB 403 TGGCAGCAGCTGCTTACGGGCTTGTCCATGCGCCGCTTGTGACAGCAGCCTT 458

RESULT 2
BX356664 1201 bp mRNA linear EST 05-MAY-2003
LOCUS BX356664 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
DEFINITION clone CSOD1015YB03 3-PRIME, mRNA sequence.
ACCESSION BX356664
VERSION BX356664
KEYWORDS BX356664.1 GI:30378083
SOURCE EST.
ORGANISM Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE
1 (bases 1 to 1201)
Li, W.B., Gruber, C., Jesse, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
CONTACT: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr
Library was constructed by life technologies, a division of
Invitrogen. Contact: Peng Liang Email: fliang@life.com URL:
http://fulllength.invitrogen.com/Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID: CSOD1015CA02NP1.

FEATURES
source
1..1201
Location/Qualifiers

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone_id="CSOD1015YB03"
/issue_type="PLACENTA COT 25-NORMALIZED"
/clone_id="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo (dT)
primer. Five prime and enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCWSPORT 6 vector. Library was normalized."

ORIGIN

Query Match
Best Local Similarity 12.9%; Pred. No. 0.48;
Matches 47; Conservative 168; Mismatches 148; Indels 0; Gaps 0;

QY 207 GTTTCGGCGCTGATTTTCGGGCGTCCACCAATGCTGATCCCTCGTTGGG 266
DB 745 STBSTSTTBTBCTTBTBSSBTBSSBTBSSBTBSSBTBSSBTBSSBTBSS 804

QY 267 CGCAGGCGCCCTGGGCGCCATCGGCTCACCATGCTGTGAATTCGCCAGCAT 326
DB 805 SBTCTSSSSSSSBTBSSTTTTSTKSBTTBSBSSSSSTBTBCTSTNTTSS 864
QY 327 CTATGGCTTTTATTCCTCCCTGCATGTGTCAAAAACCCATGCCCCGTTTATTCGT 386
DB 865 YSBSSTSTBTSSTBTSBSSBTTTBSBTTTBSSTTBSSTTBSSTTBSSTTBS 924
QY 387 TTTCGGCTTATCGAAGAACCTACGACATCTACCTCGGCGCAGCGCGCTGTGCGC 446
DB 925 BTSSBSSTTSSSSSTBSSTBSBSSSSSBTSKSSSSSSSSTBTTT 984
QY 447 GTGCGCACTTATCTCATGCAATAGCGTTTCACTCTACTGCTGATTTGGCGGCTCAG 506
DB 985 VTSSBSSTTBSSTTBSSTTSSSTBSBSSSVTSSSSSSSSTBSSSNKSSSS 1044
QY 507 CGAAGTGCATCGCAGACTGATTCCTTTTGAATTAAGGCGCTGAGTCCCTTG 566
DB 1045 SSSBSSSSSSSSSBTTTBSBTTBSSBSSSTTTTSSSSSSTBTTSTTT 1104
QY 567 CTC 569
DB 1105 TTS 1107

RESULT 3
CC684922 743 bp DNA linear GSS 19-JUN-2003
LOCUS CC684922
DEFINITION OGUAX55TH ZM 0.7 1.5 KB Zea mays genomic clone ZM8MA0397014,
genomic survey sequence.
ACCESSION CC684922
VERSION CC684922.1 GI:32089698
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoidae; Andropogoneae; Zea.
1 (bases 1 to 743)
Whitehead, C.A., Quackenbush, J., Van Aken, S., Utterback, T.,
Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,
Citek, R.W., Nunez, A., Robbins, D. and Lakey, N.
Consortium for Maize Genomics
Unpublished (2002)
Other GSSs: OGUAX55TV
CONTACT: Cathy Whitehead
TIGR

REFERENCE
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitehead@tigr.org
Seq primer: TR
Class: sheared ends.

FEATURES
source
1..743
Location/Qualifiers

/organism="Zea mays"
/mol_type="genomic DNA"
/strain="873"
/db_xref="taxon:4577"
/clone_id="ZM8MA0397014"
/clone_id="ZM 0.7 1.5 KB"
/note="Vector: pBSK-; Site 1: HindIII, 0.7-1.5 kb
methylation filtered genomic DNA library"

ORIGIN

Query Match
Best Local Similarity 53.6%; Pred. No. 13;
Matches 82; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

QY 431 CCGAGAGCTGTGCGGCTGCGACCTTATCTCAATGCAATAGCGTTTCACTCTGAGG 490
DB 475 CTTCCGCTTCCGCGCGCTTCCCTCTCCACATCTCATCTCGCATGCGCAATTCGAG 534

Db 1100 SCMBBSSKCCCKKKKBCMKCMCBKCKVCKKKKKCKKKKKCKKKCMKK 1159
QY 599 GC 600
Db 1160 BC 1161

RESULT 6
CNS0091P 925 bp DNA linear GSS 03-JUN-1999
LOCUS Drosophila melanogaster genome survey sequence TET3 end of BAC #
BACR19D16 of Rpci-98 library from Drosophila melanogaster (fruit
fly), genomic survey sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
GSS.
AL053013.1 GI:4934461
Drosophila melanogaster (fruit fly)
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 925)
Genoscope.
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage;
BP 191 91006 Evry cedex - FRANCE (E-mail: seget@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see <http://www.fruitfly.org/TheBDGP/Drosophila>
melanogaster BAC library was prepared by Kazuhiro Osoegawa and
Aaron Mammoler in Pieter de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named Rpci-98 and was constructed by partial
EcoRI digestion of Drosophila DNA provided by the BDGP from the
isogenic strain Y2; cn bw sp, the same strain used for the BDGP's
P1 and EST libraries. A more detailed description of the library
and how to order individual BAC clones, the entire library, or
filters for hybridization from the BACPAC Resource Center can be
found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES
source
1..925
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACR19D16"
/clone_lib="Rpci-98"
/note="end : TET3"

ORIGIN
Query Match 5.2%; Score 39; DB 29; Length 925;
Best Local Similarity 14.2%; Pred. No. 19;
Matches 37; Mismatches 106; Indels 0; Gaps 0;
QY 195 GCGAGCCCACTGTTTCCGGCTGATTTTGGCGGCTCCACGGAATGCTGTCATCGC 254
Db 638 SSSSSSTSSSTSSSTSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSS 697
QY 255 COTGTTGGCGGAGCCCGCTGCGCATCGCGCTCACCATTTGCTGCTGAACTT 314
Db 698 SSSSSSTSSSSSVSSSSKSTBSGSSSSSSSSSSSSSSSSSSSSSSSSSSST 757
QY 315 CCGCGACGATTTCTATGCGTTTTCATTTCCGCTGCATGTGTCAAAAACCCCATTTGCCG 374
Db 758 CACTCCCTCCSVSSSTSSSSSTSSSSSTSSSSSSSSSSSSSSSSSSSSSSSS 817
QY 375 TTTCATTCGGTTTTCGGCTTATGACGAAGCCATGACGCTGCGGCGACAGCCGCG 434
Db 818 YBMCYCTSTCGGSSSSSSGKGVTKCGCGGSSSTNMBSTSSACSSSSSSSSSSSVSS 877

QY 435 AGGCTGTCGGCGTGGCGAC 454
Db 878 SSKSSASSSSSVSSSSGSSGS 897

RESULT 7
BX361080 1201 bp mRNA linear EST 05-MAY-2003
LOCUS BX361080 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
DEFINITION clone CSOD1079YN16 3-PRIME, mRNA sequence.
BX361080
ACCESSION BX361080
VERSION BX361080.1 GI:30374504
KEYWORDS EST.
SOURCE
ORGANISM
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 1201)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
CONTACT: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seget@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen, contact : Feng Liang Email : liang@lifetech.com URL :
<http://fulllength.invitrogen.com/InvitrogenCorporation1600>
Paraday Avenue Genoscope sequence ID : CSOD1079DG08NP1.

FEATURES
source
1..1201
/organism="Homo sapiens"
/mol_type="RNA"
/db_xref="taxon:9606"
/clone="CSOD1079YN16"
/clone_lib="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-clig0 (dt)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN
Query Match 5.1%; Score 38.6; DB 13; Length 1201;
Best Local Similarity 10.5%; Pred. No. 28;
Matches 34; Mismatches 142; Indels 0; Gaps 0;
QY 19 ATTCATTCAAGCTGAGAGTGTGCGCATCAAGCGAGCGCTGAACGATGATTAAGT 78
Db 825 KDDMMMMKKNDKK 884
QY 79 TATCGCGCTACGAATGCGCAAGTCTAAAACTCTCTGTCGAGTTTGGGCTG 138
Db 885 KKKKKDAKKKKWAKKKKKKAAKKKKKKKKKKKKKKKKKKKKKKKKKKKKKK 944
QY 139 TACCGATTGTGATTCGCTGCTCTGTTATTAATACGGTACGAATGCGGCA 198
Db 945 KKKKKKKKKKKTTTWTTTTWTMTMMATTTTWTTTTTHKKKKKKKKKK 1004
QY 199 GCCCACTGTTTTCGGCTGATTTTTCGGGCTCCACGGAATGCTGATCGCCCTC 258
Db 1005 KNDMMMMKK 1064
QY 259 GTTGGGCGGAGCCCGCTGCGCATCGCGCTCACACATTTCTGTGAATCTTCCG 318
Db 1065 KKKKKKKCCCBKCCCKKCKBKCYCKKKKKKKKKKKKKKKKKKKKKKKKK 374
QY 319 CAGTATTCATGCGTTTTCATT 341
Db 1125 CCKKKKKKCKKKKKKKKKKK 1147

RESULT 8

CB657749 835 bp mRNA linear EST 09-APR-2003
 LOCUS OSJNEC1F02.f OSJNEC Oryza sativa (japonica cultivar-group) cDNA
 DEFINITION clone OSJNEC1F02 5', mRNA sequence.
 ACCESSION CB657749
 VERSION CB657749.1 GI:29661474
 KEYWORDS EST.
 SOURCE Oryza sativa (japonica cultivar-group)
 ORGANISM Oryza sativa (japonica cultivar-group)
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Ehrhartoideae; Oryzaceae; Oryza.
 1 (bases 1 to 835)
 Jantanasuriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,E.,
 Kudirna,D., Dean,R., Soderlund,C., Wing,R. and Wang,G.
 Large-scale identification of ESTs involved in the interaction
 between rice and Magnaporthe grisea
 Unpublished (2003)
 JOURNAL Contact: Rod Wing
 COMMENT Arizona Genomics Institute
 University of Arizona
 Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
 85721-0088, USA
 Tel: 520 626 3967
 Fax: 520 621 9288
 Email: http://genome.arizona.edu
 PCR Primers
 FORWARD: gta aac cga cgg cca gtc
 BACKWARD: gga aac agc tat gac cat g
 Plate: 13 row: F column: 02
 Seq primer: gta aac cga cgg cca gtc.
 Location/Qualifiers
 1..835
 /organism="Oryza sativa (japonica cultivar-group)"
 /mol_type="mRNA"
 /cultivar="Nipponbare"
 /db_xref="taxon:39947"
 /clone="OSJNEC1F02"
 /tissue_type="leaf"
 /dev_stage="3 week"
 /lab_host="DH10B"
 /clone_lib="OSJNEC"
 /note="Vector: pBluescript II KS +; Site 1: EcoRI; Site 2:
 XhoI; 6 hrs after inoculation with Rice Blast (C9240-1)"

ORIGIN
 Query Match 5.0%; Score 38; DB 14; Length 835;
 Best Local Similarity 51.1%; Pred. No. 34;
 Matches 89; Conservative 0; Mismatches 85; Indels 0; Gaps 0;
 252 CGCCCTGCTGTGAGGCGAGCGCCCTGAGCGGCATCGGCTCACCACATTTGCTGTGAA 311
 79 CGACCGGCTCTCGAGGCGCGCTGCTCGGCGCTTCGCGCGCTGCGAGATGCCCTC 138
 312 CTTCGCCGACGATTTCTATGCGTTTCATTCCTCCGCTGCATGTGCTAATAAACCCTATTC 371
 139 ATCCCGCGGGGTGTGACCGGATCGGACCCCGACCTGCGCGCTGGAACGCGCTGCT 198
 372 CCGTTTCATTCGAGTTTTCGCGCTTATCGACGAGCTTACGACGAGTCACTGCGGC 425
 199 CTCGGCTACGCGCGCTTGTGTCCAGGACGTCGTGCGCTGCGACCTCTGCGGC 252

RESULT 9
 BH887860 697 bp DNA linear GSS 07-AUG-2002
 LOCUS LB01712a.d SP6.1 Leishmania major Friedlin BAC Library Leishmania
 DEFINITION major genomic clone LB01712a, genomic survey sequence.
 ACCESSION BH887860
 VERSION BH887860.1 GI:22133383
 KEYWORDS GSS.
 SOURCE Leishmania major
 ORGANISM Leishmania major

Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae;
 Leishmania.
 1 (bases 1 to 697)
 Myler,P.J., Vogt,C., Munden,H., Robertson,L., Slak,E.,
 Ravelin,A.G., Aggarwal,G., Nelson,S., Seyler,A., Worthey,E.,
 Stuart,K. and Ragland,M.
 Leishmania major Friedlin BAC End Sequences
 Unpublished (2002)
 Other GSSs: LB01712a.d.T7.1
 JOURNAL Contact: Myler PJ
 COMMENT Seattle Biomedical Research Institute
 4 Nickerson Street, Seattle, WA 98109-1651, USA
 Tel: 206 284-8846
 Fax: 206 284-0313
 Email: mylerp@bri.org
 Seq primer: SP6
 Class: BAC ends.
 Location/Qualifiers
 1..697
 /organism="Leishmania major"
 /mol_type="genomic DNA"
 /strain="Friedlin"
 /db_xref="taxon:5664"
 /clone="LB01712a"
 /lab_host="E. coli GeneHog + TrfA"
 /clone_lib="Leishmania major Friedlin BAC Library"
 /note="Vector: pCG270; Site 1: HindIII; Genomic DNA from
 Leishmania major Friedlin in agarose blocks was partially
 digested with HindIII, size selected, and ligated with
 HindIII-digested pCG270 vector DNA. 10368 clones were
 picked and arrayed in 384- and 96-well plates. Library
 construction and arraying was carried out by Resgen
 Corporation and clones and filters are available from
 them"

ORIGIN
 Query Match 5.0%; Score 37.4; DB 28; Length 697;
 Best Local Similarity 52.2%; Pred. No. 46;
 Matches 83; Conservative 0; Mismatches 76; Indels 0; Gaps 0;
 181 GCTACGAAATGTTGGGAGACCCCACTGTTTCGCGCTGATTTTCGGGCTCAACGAA 240
 579 GCGAGCGATGCTGTGTGTCGCGACTCTCCACCCCAACGATGCTCGCGCGCG 520
 241 ATGCTGTCATCGCCCTGTTGTGGGCGACGCGCCCTGAGCGGCATCGGCTCACCACA 300
 519 CACGAGTCGGCGCGCTGCTGCGAGCCTCGTCCGCTTTCGGGCAACCGTTACCAAG 460
 301 TTGCTGTGAACCTTCGCGACGATTTCTATGCGCTTTCA 339
 459 CCACTGTCACTGTGATCATGCTCCGTGCGCATTCGA 421

RESULT 10
 BX382084/c 1108 bp mRNA linear EST 08-MAY-2003
 LOCUS BX382084 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
 DEFINITION clone CS0D1074YC24 5-PRIME, mRNA sequence.
 ACCESSION BX382084
 VERSION BX382084.1 GI:30455015
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 1108)
 Li,W.B., Gruber,C., Jessee,J. and Pelayes,D.
 Full-length cDNA libraries and normalization
 Unpublished (2001)
 JOURNAL Contact: Genoscope
 COMMENT Genoscope - Centre National de Sequencage
 BP 191 9106 Evry cedex - France
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr

Library was constructed by Life Technologies, a division of Invitrogen. Contact : Feng Liang Email : fliang@lifetech.com URL : <http://fulllength.invitrogen.com/InvitrogenCorporation1600>
 Faraday Avenue Genoscope sequence ID : CS0D1074BB120P1.

FEATURES

source

1. 1108
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0D1074YC24"
 /issue_type="PLACENTA COT 25-NORMALIZED"
 /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
 /note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 5.0%; Score 37.4; DB 13; Length 1108;
 Best Local Similarity 16.3%; Pred. No. 57;
 Matches 65; Conservative 145; Mismatches 189; Indels 0; Gaps 0;

114 CTCCTGCTGCGAGGTTGGGATGATACCGGATGATGCTGCTGCTGAT 173
 540 CTTTSTKSTKTTGTTTACCTTAMCCTKCTKTKACGGGCTKTCWCK 481
 174 TCAATACGGCTACGAATGTGGGAGCCCATGTTTCCGGCTGATTTCCGGGCTC 233
 480 TSTTCTKCTDCKGCGCTKTKTKDCTTKKCMCKMKCCCKSTKCTKTKTKKST 421
 234 CACCGAATGCTGATGCGCCCTCGTTGGGCGGAGCGGCCCTGGGCGCATCGCT 293
 420 WCKTCTTTTCKKCKCKCTKSKTKTKWCKSKCKSKCKSKCKSKCKSKCKSKCK 361
 294 CACCAATGCTGATGCGCCCTCGTTGGGCGGAGCGGCCCTGGGCGCATCGCT 353
 360 SSACKSTKCKCTKSKSKCKSKCKSKCKSKCKSKCKSKCKSKCKSKCKSKCK 301
 354 GGTCAAAAACCCATGCGCGTTTCTATGCGTTTCCGGCTGATTTCCGGGCTC 413
 300 TSTTATTSATACGAMTKSSMSMSMKMTGTTMKSKDOTTSGMKASCCKTKSS 241
 414 AGTCACTGCGCGAGCGCCGAGCTGTGCGGCGGCGGACCTATGTCGAATAC 473
 240 KSSATATGKMSCTGSMTCSSCKSSCKSSCKSSCKSSCKSSCKSSCKSSCK 181
 474 GTTCACTCTCTACCTGCGGATTCGGCGCTCTCACCGAGT 512
 180 TSMTKMTKMSSTTKSKKMKMKKAKSKSATSKKK 142

RESULT 11
 BG845919/c 730 bp mRNA linear EST 29-MAY-2001
 LOCUS 1024011F06.x1 C. reinhardtii CC-1690, normalized, Lambda Zap II
 DEFINITION Chlamydomonas reinhardtii cDNA, mRNA sequence.
 ACCESSION BG845919
 VERSION BG845919.1 GI:14227103
 KEYWORDS EST.
 SOURCE Chlamydomonas reinhardtii
 ORGANISM Chlamydomonas reinhardtii
 Chlamydomonas reinhardtii
 Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
 Chlamydomonadales; Chlamydomonas.
 1 (bases 1 to 730)
 Grossman, A., Davies, J., Federpiel, N., Harris, E., Lefebvre, P.,
 McInerney, J. P., Silflow, C., Stern, D. and Surzycki, R.
 Analysis of the Chlamydomonas reinhardtii Genome: A Model,
 Unicellular System for Analyzing Gene Function and Regulation in
 Vacuolar Plants project phase 2
 Unpublished (2000)
 JOURNAL Contact: Charles Hauser
 COMMENT DCMB Box 91000
 Duke University

Durham, NC 27708-1000
 Tel: 919 613 8159
 Fax: 919 613 8177
 Email: chauser@duke.edu
 Location/Qualifiers

FEATURES

source

1. 730
 /organism="Chlamydomonas reinhardtii"
 /mol_type="mRNA"
 /strain="CC-1690 wild type mt+ 21gr"
 /db_xref="taxon:3055"
 /clone_lib="C. reinhardtii CC-1690, normalized, Lambda Zap II"
 /note="Vector: Bluescript II SK-; Site 1: EcoRI; Site 2: XhoI. This library, constructed by John Davies and Jeffrey McInerney, combines cDNAs from CC-1690 cells grown to mid-log phase in YAP (acetate-containing) medium in the light, YAP medium in the dark, HS (minimal) medium in ambient levels of CO2 and HS medium bubbled with 5% CO2. PolyA mRNA was purified from each sample, pooled and cDNA synthesized. The cDNA was directionally cloned into lambda Zap II (Stratagene) in the EcoRI (5') and XhoI (3') sites. Bluescript II SK- plasmids were excised from the lambda Zap clones by superinfection with ExAssist (Stratagene) phage. The library was normalized using method 4 described in Bonaldo et al (1996) Genome Research 6: 791-806."

ORIGIN

Query Match 4.9%; Score 37.2; DB 12; Length 730;
 Best Local Similarity 49.5%; Pred. No. 53;
 Matches 96; Conservative 0; Mismatches 98; Indels 0; Gaps 0;

202 CCACGTTTCCGCGCTGATTTTCGCGGCTCCACCGAAATGCTGATCGCCCTGTT 261
 631 CCCGGTGTCAAGGCGCGGGAAGCCGCGCCCAAGATTACGAGGATCGGCGCGGC 632
 262 GTGGCGGAGCGCCCTGGGCGCATGCGCTACCACTTCTGCTGATCGCGCAC 321
 631 TTGTGGCGCGGCTGTGAGCCCGCGTATCGACGAGGTGGCGAGTGTCCAGTAC 572
 322 GTATTCATGCGCTTTTATTCCTCCGCTGATGCTGATCAAAAACCCATGCGCCCTTTAT 381
 571 GACGCAATCGACATGAGCTCCCGCGTGGAGAGAGGCGCTCATGTGGGATGAGC 512
 382 TCGGTTTTCGCGCT 395
 511 TCGGCGCGCGCGGT 498

RESULT 12
 BF101158/c 1454 bp mRNA linear EST 19-OCT-2000
 LOCUS 601754733F1 NCI_CGAP_Mami Mus musculus cDNA clone IMAGE:3963528 5',
 DEFINITION mRNA sequence.
 ACCESSION BF101158
 VERSION BF101158.1 GI:10883684
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NIH-MGC <http://mgi.nci.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 CONTACT: Robert Strussberg, Ph.D.
 Email: cgasbe-remail.nih.gov
 Tissue Procurement: Gilbert Smith, Ph.D.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>

II;

SECRET

Tue Apr 20 06:47:17 2004

us-10-608-504-2.rst

Page 8

ORIGIN /plasmid="pbeloBAC11"
/note="end : T7"

Query Match 4.9%; Score 36.8; DB 29; Length 1201;
Best Local Similarity 20.4%; Pred. No. 87;
Matches 73; Conservative 131; Mismatches 153; Indels 1; Gaps 1;

QY 38 TGTGGCCATCCAGAGCGCCCTGGAACAGATGATTAAGTTATCGCCCTACGAATCG 97
DB 823 TTTCCTCCAMCMCMAMMMCCCTTCMAGTANCTCMAMMMMGSGGCAAMVVMV 882
QY 98 CGCAGAGCTTAAAAACCTCCCTTGTGCGAGGTTGGGCGATGTACCCGATTGTATGCGT 157
DB 883 VVMVMSYCCCMAGMACCHTCAATKCCMGDASCCGDSVSSSGSCCVRSSTVGRMYH 942
QY 158 TTGATCTCTGTGTTATTCATACGAGCTACGATGTAGGAGC-CCCACTGTTTCCGCG 216
DB 943 YTGCGACCCGCTTMTWSCSSSGSSGCCKTIVGSCSSSCSSSGSCSCBCTGSSSSSS 1002
QY 217 CTGATTTTCGGGGCTCCACCGAATGCTGTATCGCCCTGTTGTGGCGAGCGCC 276
DB 1003 SSSCBSTGTCTYSSSSSSSSSKCSKTSYKGTCTSYGCKTGKGGCCCGGSGCYC 1062
QY 277 CTGGGGCCGACGCGCTCACCAATGCTGTGTAACCTCGCCACGATTCATGAGTTT 336
DB 1063 CBCCCCGCKCCGSGSTSKSTGKTGKSSSGGGGSSSSSBSBNTSSS 1122
QY 337 TCATTCGCGCTGATGTGTAACCAACCCATGCGCCGTTTCATTCGATTCGCGC 394
DB 1123 TTTBSSSSSTSTSSGSGSBRTBSSBSSTSTBTBTBTBTBTBTBTBTBTBTBT 1180

RESULT 15
BE013856/c 539 bp mRNA linear EST 09-JUL-2000
LOCUS
DEFINITION 125354 MARC 1PIG Sus scrofa cDNA 5', mRNA sequence.
ACCESSION BE013856
VERSION BE013856.1 GI:8274843
KEYWORDS
SOURCE Sus scrofa (pig)
ORGANISM Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.
1 (bases 1 to 539)
Fahnenkrug,S.C., Smith,T.P.L., Preking,B.A., Cho,J., White,J.,
Vallet,J., Wise,T., Rohrer,G.A., Pertea,G., Sultana,R.,
Quackenbush,J. and Keefe,J.W.
Porcine gene discovery by normalized cDNA-library sequencing and
EST cluster assembly
Mamm. Genome 13 (8), 475-478 (2002)
22213789
12226715

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
COMMENT
Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called and alt trimmed with phred
v0.980904.e. Vector identified by cross_match with the -minscore 18
and -timematch 12 options.
PCR primers
FORWARD: AGGAAACAGCTATGACAT
BACKWARD: GTTTCACAGTACGACG
Plate: 52 row: G column: 21
Seq primer: ATTAGTGACACTATAG.
Location/Qualifiers
1..539
/organism="Sus scrofa"
/mol_type="mRNA"
/db_xref="taxon:9823"
/tissue_type="pooled"

ORIGIN /lab_host="DH10B"
/clone_id="MARC 1PIG"
/note="Vector: PCMV SPOR6; Site_1: NotI; Site_2: SalI;
Library made from pooled tissue from day 11, 13, 15, 20,
and 30 embryos."

Query Match 4.9%; Score 36.6; DB 10; Length 539;
Best Local Similarity 54.0%; Pred. No. 67;
Matches 75; Conservative 0; Mismatches 64; Indels 0; Gaps 0;

QY 211 TCCGACCTGATTTTCGGGGGCTCCACCGAATGCTGCTCATCGCCCTGTTGTGGCGCA 270
DB 445 TCTGCTGGGCTTCCCGAGGCGCCAGCTGAGGGGTGTGAGGGCTGTACAGATCA 386
QY 271 GCGCCCTGGGCGGCGCTACCAATGCTGTGTAATTCCGCCACATATTCTAT 330
DB 385 CCGGCCAGGAGAGCAGCACCCGCCACCTCTCTGACAGCCAGGAGCGGAGTCAAG 326
QY 331 GCGTTTCATTCGCGCTGC 349
DB 325 GTGTCTGACAGCGGCTGC 307

Search completed: April 17, 2004, 20:01:43
Job time : 2440.38 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 16, 2004, 06:51:12 ; Search time 70.6156 Seconds
(without alignments)
1004.303 Million cell updates/sec

Title: US-10-608-504-3
Perfect score: 1269
Sequence: 1 MORTORINHSLEVSPEKAL.....LIFGLTLTRYFLGKAK 251

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues
Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database :
1: Genesecp29Jan04:*
2: genesecp1990s:*
3: genesecp2000s:*
4: genesecp2001s:*
5: genesecp2002s:*
6: genesecp2003as:*
7: genesecp2003bs:*
8: genesecp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1269	100.0	251	4 AAG90035	Aag90035 C. glutam.
2	1265	99.7	251	4 AAB86247	Aab86247 C. glutam.
3	428.5	33.8	237	4 AAG31154	Aag31154 C. glutam.
4	214	16.9	338	5 ABP66261	Abp66261 Bifidobac.
5	197.5	15.6	261	6 ABM69066	Abm69066 Photobac.
6	197	15.5	245	5 ABP57756	Abp57756 E. coli L
7	190	15.0	266	6 ADA35238	Ada35238 Actinoba.
8	189.5	14.9	263	6 ADA33199	Ada33199 Actinoba.
9	185.5	14.6	255	7 ADC95293	Adc95293 E. faeciu
10	177.5	14.0	218	6 ABU00481	Abu00481 S. pneumo
11	177.5	14.0	218	6 ABP81300	Abp81300 Streptoco
12	174.5	13.8	235	5 ABM54958	Abm54958 Lactococ
13	169	13.3	235	5 ABM47959	Abm47959 Listeria
14	167	13.2	230	5 ABP30530	Abp30530 Streptoco
15	167	13.2	231	5 ABP28562	Abp28562 Streptoco
16	161.5	12.7	242	5 ABP40309	Abp40309 Staphyloc
17	158.5	12.5	230	4 AAG81587	Aag81587 S. epider
18	150.5	11.9	226	6 ADB08660	Adb08660 Allotiococ
19	148.5	11.7	214	6 ADB08658	Adb08658 Allotiococ
20	148	11.7	228	6 ABM72815	Abm72815 Staphyloc
21	141.5	11.2	247	6 ADA34737	Ada34737 Actinoba
22	114	9.0	475	4 AAU64076	Aau64076 Propionib
23	114	9.0	475	6 ABM60595	Abm60595 Propionib
24	112.5	8.9	876	5 ABP55229	Abp55229 Bifidobac
25	112	8.8	379	5 ABB55268	Abb55268 Lactococ

26	111	8.7	113	3 AAY74753	Aay74753 Neisseria
27	105	8.3	417	7 ADD46525	Add46525 Human pro
28	105	8.3	417	7 ADD46529	Add46529 Human pro
29	104	8.2	377	6 ADB12483	Adb12483 Allotiococ
30	104	8.2	414	6 ADB12481	Adb12481 Allotiococ
31	104	8.2	442	6 ADB12479	Adb12479 Allotiococ
32	102	8.0	112	3 AAY74752	Aay74752 Neisseria
33	101.5	8.0	746	7 ADB75242	Adb75242 Prostata
34	101.5	8.0	766	4 AAM79259	Aam79259 Human pro
35	101.5	8.0	816	6 ADA55075	Ada55075 Human pro
36	101.5	8.0	820	5 ABP56272	Abp56272 Human TCH
37	101.5	8.0	1203	4 ABG05471	Abg05471 Novel hum
38	101.5	8.0	1597	4 ABG19902	Abg19902 Novel hum
39	101.5	8.0	1597	4 ABG10812	Abg10812 Novel hum
40	101.5	8.0	1597	4 ABG10253	Abg10253 Novel hum
41	101.5	8.0	1597	4 ABG03989	Abg03989 Novel hum
42	101.5	8.0	1599	4 ABG07884	Abg07884 Novel hum
43	101.5	8.0	1599	4 ABG14450	Abg14450 Novel hum
44	101.5	8.0	1784	4 ABG09148	Abg09148 Novel hum
45	98.5	7.8	478	3 AAG15416	Aag15416 Arabidops

ALIGNMENTS

RESULT 1
AAG90035
ID AAG90035 standard; protein; 251 AA.
AC AAG90035;
DT 26-SEP-2001 (first entry)
DE C glutamicum protein fragment SEQ ID NO: 3789.
KW Corynebacterium; amino acid synthesis; vitamin; saccharide;
KW organic acid synthesis.
OS Corynebacterium glutamicum.
PN EP1108790-A2.
PD 20-JUN-2001.
PP 18-DEC-2000; 2000EP-00127668.
PR 16-DEC-1999; 99JP-00377484.
PR 07-APR-2000; 2000JP-00159162.
PR 03-AUG-2000; 2000JP-00280988.
PA (KYOW) KYOWA HAKKO KOGYO KK.
PI Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;
PI Tateishi N, Senoh A, Ikeda W, Ozaki A;
DR WPI; 2001-376931/40.
DR N-PSDB; AAH65254.
PT Novel polynucleotides derived from Corynebacterium, for identifying
PT mutation point of a gene, measuring expression of a gene, analyzing
PT expression profile or pattern of a gene and identifying homologous gene.
PS Claim 17; SEQ ID NO 3789; 245pp + Sequence Listing; English.
The present invention provides a number of nucleotide and protein
sequences from the Corynebacterium glutamicum. These
are useful for identifying the mutation point of a gene derived from a
mutant of Corynebacterium, measuring expression amount and analyzing
the expression profile or expression pattern of a gene derived from
Corynebacterium, and identifying a homologue of a gene derived from
Corynebacterium. Corynebacterium are useful for producing amino
acids, nucleic acids, vitamins, saccharides and organic acids,
particularly L-lysine. The present sequence is a protein described in the

CC exemplification of the invention. Note: The sequence data for this patent
CC did not form part of the printed specification, but was obtained in
CC electronic format directly from the European Patent Office

XX Sequence 251 AA;

Query Match 100.0%; Score 1269; DB 4; Length 251;
Best Local Similarity 100.0%; Pred. No. 3.3e-137;
Matches 251; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MOKTQEHSSLEVSPPSKALEPDDKGYRREIAQGLKTSIAAGIGWYPIGIAFGILVIOY 60
DB 1 MOKTQEHSSLEVSPPSKALEPDDKGYRREIAQGLKTSIAAGIGWYPIGIAFGILVIOY 60
QY 61 GYEWMAAPLPSGLIFAGSTEMVIALVGAAPLGAIALTTLLVFRHVFYAFSPFLHYVK 120
DB 61 GYEWMAAPLPSGLIFAGSTEMVIALVGAAPLGAIALTTLLVFRHVFYAFSPFLHYVK 120
QY 121 NPIARFYSVFALIDEAVAVTARPAAGSAMRLISQIAFHSYWFVGGITGVAIAELIPE 180
DB 121 NPIARFYSVFALIDEAVAVTARPAAGSAMRLISQIAFHSYWFVGGITGVAIAELIPE 180
QY 181 IKGLEPALSCLFVTLTLDSCRTKQIPSLLAGSFTTALVVIPOALFALLIFLGILLT 240
DB 181 IKGLEPALSCLFVTLTLDSCRTKQIPSLLAGSFTTALVVIPOALFALLIFLGILLT 240
QY 241 IRYFFLGKAAK 251
DB 241 IRYFFLGKAAK 251

RESULT 2

ID AAB86247 standard; protein; 251 AA.

AC AAB86247;

DT 05-SEP-2001 (first entry)

DE C. glutamicum bnf protein.

KM L-amino acid production; bnf; bnf; branched-chain amino acid;

KM coryneform bacterium; leucine; isoleucine; valine; medicine;

OS animal nutrition.

OS Corynebacterium glutamicum.

PN EP1096010-A1.

PD 02-MAY-2001.

PF 11-OCT-2000; 2000EP-00122057.

PR 27-OCT-1999; 99DB-01051708.

PA (DEGS) DEGUSA AG.

PA (KERU) FORSCHUNGSZENTRUM JUELICH GMBH.

PI Kernerkecht N, Eggeling L, Sahn H, Pfeffeler W;

DR WPI; 2001-391595/42.

DR N-PSDB; AAH21109, AAH21110.

XX New export genes from coryneform bacteria, useful for increasing

PT fermentative production of branched-chain amino acids.

PS Claim 1a; Page 15-16; 23pp; German.

CC This invention describes a novel isolated polynucleotide (i) containing
CC at least one sequence that (i) is 70% identical with a sequence that
CC encodes at least one of 251 amino acid (aa) (3) or 108 aa (5)
CC polypeptides; (ii) encodes a polypeptide at least 70% identical with (3)
CC or (5); (iii) is the complement of (i) or (ii), or (iv) contains at least

CC 15 consecutive bases from (i)-(iii). The invention also describes (a)
CC protein derived from sequences of 753 bp (2) or 324 bp (4); (b)
CC coryneform microorganisms, especially Corynebacterium, transformed with
CC one or more (i), where these are replicative DNA; (c) production of
CC branched-chain L-aa by fermentation of coryneform bacteria in which the
CC bnf and/or bnf genes (or equivalent sequences) are amplified,
CC especially overexpressed; and (d) method for isolating the bnf and/or
CC bnf genes. (1) is used for transformation of coryneform bacteria being
CC used for fermentative production of branched-chain amino acids;
CC specifically leucine, isoleucine and valine, which are useful in medicine
CC and animal nutrition. (1) can also be used as source of primers and
CC probes for isolation of related sequences. Transformation with (i)
CC increases yield of branched-chain amino acids. This sequence represents
CC the Corynebacterium glutamicum ATCC 14752 bnf protein described in the
CC method of the invention

XX Sequence 251 AA;

Query Match 99.7%; Score 1265; DB 4; Length 251;
Best Local Similarity 99.6%; Pred. No. 9.4e-137;
Matches 250; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MOKTQEHSSLEVSPPSKALEPDDKGYRREIAQGLKTSIAAGIGWYPIGIAFGILVIOY 60
DB 1 MOKTQEHSSLEVSPPSKALEPDDKGYRREIAQGLKTSIAAGIGWYPIGIAFGILVIOY 60
QY 61 GYEWMAAPLPSGLIFAGSTEMVIALVGAAPLGAIALTTLLVFRHVFYAFSPFLHYVK 120
DB 61 GYEWMAAPLPSGLIFAGSTEMVIALVGAAPLGAIALTTLLVFRHVFYAFSPFLHYVK 120
QY 121 NPIARFYSVFALIDEAVAVTARPAAGSAMRLISQIAFHSYWFVGGITGVAIAELIPE 180
DB 121 NPIARFYSVFALIDEAVAVTARPAAGSAMRLISQIAFHSYWFVGGITGVAIAELIPE 180
QY 181 IKGLEPALSCLFVTLTLDSCRTKQIPSLLAGSFTTALVVIPOALFALLIFLGILLT 240
DB 181 IKGLEPALSCLFVTLTLDSCRTKQIPSLLAGSFTTALVVIPOALFALLIFLGILLT 240
QY 241 IRYFFLGKAAK 251
DB 241 IRYFFLGKAAK 251

RESULT 3

ID AAG93154 standard; protein; 237 AA.

AC AAG93154;

DT 26-SEP-2001 (first entry)

DE C glutamicum protein fragment SEQ ID NO: 6908.

KM Coryneform bacterium; amino acid synthesis; vitamin; saccharide;

KM organic acid synthesis.

OS Corynebacterium glutamicum.

PN EP108790-A2.

PD 20-JUN-2001.

PF 18-DEC-2000; 2000EP-00127688.

PR 16-DEC-1999; 99JP-00377484.

PR 07-APR-2000; 2000JP-00159162.

PR 03-AUG-2000; 2000JP-00280988.

PA (KYOW) KYOWA HAKKO KOGYO KK.

PI Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;

PI Tateishi N, Senoh A, Ikeda M, Ozaki A;

DR WPI, 2001-376931/40.
DR N-PSDB; AAH68373.
XX Novel polynucleotides derived from *Corynebacterium* bacteria, for identifying
PT mutation point of a gene, measuring expression of a gene, analyzing
PT expression profile or pattern of a gene and identifying homologous gene.
XX
PS Claim 17; SEQ ID NO 6908; 246bp + Sequence Listing; English.
XX
CC The present invention provides a number of nucleotide and protein
CC sequences from the *Corynebacterium* *Corynebacterium glutamicum*. These
CC are useful for identifying the mutation point of a gene derived from a
CC mutant of *Corynebacterium* bacterium, measuring expression amount and analysing
CC the expression profile or expression pattern of a gene derived from
CC *Corynebacterium* bacterium, and identifying a homologous of a gene derived from
CC *Corynebacterium* bacterium. *Corynebacterium* bacteria are useful for producing amino
CC acids, nucleic acids, vitamins, saccharides and organic acids,
CC particularly L-lysine. The present sequence is a protein described in the
CC exemplification of the invention. Note: The sequence data for this patent
CC did not form part of the printed specification, but was obtained in
CC electronic format directly from the European Patent Office
XX
SQ Sequence 237 AA;
XX
Query Match 33.8%; Score 428.5; DB 4; Length 237;
Best Local Similarity 40.1%; Pred. No. 1.5e-40;
Matches 87; Conservative 46; Mismatches 81; Indels 3; Gaps 3;
XX
QY 31 ETGCGKTSIAGLGMVPIPIGIAAGLVITQGYEMWAPLPSGILPFGSTEMLVITLV-VG 89
DB 8 EIRGGIGETLVGLGIPGLAAGLVITQGYEMWAPLPSGILPFGSTEMLVITLV-VG 67
QY 90 AAPLGIATLTLLNPRHVFYAFSPPLHYVKNPIARFYGVFALIDRAVVTARAPAG-WS 148
DB 68 IGFESA-AVAGFVWNRHIFGILTPFRHRIKSGAGRAYTALYTDSSYAIYVGRPPGDS 126
QY 149 AMRLIMQIAFHSYVYFGILTVGVAIAELIPPEIKGLEFALCSLFTVTLTDSCKTKQIPSS 208
DB 127 GTRVLTVQILCOALMWITPGITGALVGVVPPDKKMDFFALTLFVLAWEAKNNKDYSL 186
QY 209 LILAGLSTIALVVIPOALFALLIFLGLITRYPF 245
DB 187 PLFAVVLAVVSGFVAEOMLVITLTYFLLILRYVF 223
XX
RESULT 4
ABP6261
ID ABP6261 standard; protein; 338 AA.
XX
AC ABP6261;
XX
DT 19-NOV-2002 (first entry)
XX
DE Bifidobacterium longum NCC2705 ORF amino acid sequence SEQ ID NO:1005.
XX
KM Bifidobacterium longum NCC2705; Bifidobacterium; bacterial;
KM antiarrhythmic; antibacterial; inhibitor of *Salmonella*; detection;
KM identification; lactic acid bacterium; diarrhoea; pathogenic bacteria;
KM rotavirus; food composition; pharmaceutical composition.
XX
OS Bifidobacterium longum.
XX
PN EP1227152-A1.
XX
PD 31-JUL-2002.
XX
PF 30-JAN-2001; 2001EP-00102050.
XX
PR 30-JAN-2001; 2001EP-00102050.
XX
PA (NEST) SOC PROD NESTLE SA.
XX
DR WPI, 2002-668397/72.

XX
PT Novel polynucleotide comprising Bifidobacterium genome sequence useful as
PT a probe or primer for detecting and/or identifying Bifidobacterium longum
PT in a biological sample.
XX
PS Claim 3; SEQ ID NO 1005; 80bp; English.
XX
CC The present invention describes a polynucleotide (I) comprising a
CC sequence of a Bifidobacterium genome selected from the nucleotide
CC sequences given in AB081842 and AB081843, or a sequence exhibiting at
CC least 90% identity or which hybridises with the sequences given in
CC AB081842 and AB081843. Also described is a polynucleotide (II) encoding a
CC fusion protein, comprising a sequence selected from 1097 sequences given
CC in ABP65288 to ABP63351 ligated in frame to a polynucleotide encoding a
CC heterologous polypeptide. (II) has antiarrhythmic and antibacterial
CC activities, and can be used as an inhibitor of *Salmonella*. (I) (which is
CC a probe) is useful for the detection and/or identification of
CC Bifidobacterium longum in a biological sample. A carrier containing the
CC lactic acid bacterium Bifidobacterium longum NCC2705 (CNCM I-2618) can be
CC used for preventing and/or treating diarrhoea brought about by pathogenic
CC bacteria and/or rotavirus. The carrier is a food composition selected
CC from milk, yogurt, curd, cheese, fermented milks, milk based fermented
CC products, ice-creams, fermented cereal based products, milk based
CC powders, infant formula, pet food or a pharmaceutical composition
CC selected from tablets, liquid bacterial suspensions, dried oral
CC supplement, wet oral supplement, dry tube feeding or wet tube feeding.
CC (I) is useful in DNA arrays or chips to carry out analysis of the
CC expression of the Bifidobacterium gene. AB081844 to AB081850 represent
CC Bifidobacterium related nucleotide sequences given in the Sequence
CC Listing from the present invention but not mentioned further within the
CC specification. N.B. The sequence data for this patent is not represented
CC in the printed specification but is based on sequence information
CC supplied by the European Patent Office
XX
SQ Sequence 338 AA;
XX
Query Match 16.9%; Score 214; DB 5; Length 338;
Best Local Similarity 29.9%; Pred. No. 1.1e-15;
Matches 64; Conservative 37; Mismatches 85; Indels 28; Gaps 8;
XX
QY 44 LGMPDGIAGFLVLYOYGEWMAAPLPSGILPFGSTEMLVITLVGVA-APLGIATLTLL 102
DB 99 LGPLFLGASVILGNTGKFSFVWPMCSAPTFRQSMFVYVNLISFNLAGE-LIALM 157
QY 103 VNRHVFYAFSPPLHYVKNPIARF-----YVFALIDEAYAV--TAARPAGWS-AMR 151
DB 158 VNARHLFYGISM-----LGKFKGLGWRKPYLIPGMCDETFAINSTAKIPAGIDRGWF 209
QY 152 LISMQIAFHSYVYFGILTVGVAIAELIPPEIKGLEFALCSLFTVTLTDSCKTKQIPSL-L 210
DB 210 YFWVTLQNLQVWVGTALGILGHLHPNTDGLDFVLTALFLVFLDQMDGKRERLSA 269
QY 211 LAGLSTIALVVIPOALFALLIFLGLITRYPF 245
DB 270 VIQVLTSLACLLIFGANDFMIPSMIAMLLIFVAL 303
XX
RESULT 5
ABM69066
ID ABM69066 standard; protein; 261 AA.
XX
AC ABM69066;
XX
DT 20-NOV-2003 (first entry)
XX
DE Photorhabdus luminescens protein sequence #2163.
XX
KM Antibacterial; fungicide; insecticide; polymorphism; genetic analysis;
KM detection; food; gene expression; plant; animal; microorganism; toxin;
KM antibiotic; biopesticide; virulence factor; disease model; plague;
KM whooping cough.
XX
OS Photorhabdus luminescens.

XX WO200294867-A2.
 XX 28-NOV-2002.
 PD 07-FEB-2002; 2002WO-IB003040.
 XX 07-FEB-2001; 2001FR-00001659.
 XX 07-FEB-2001; 2001FR-00001659.
 XX (INSP) INST PASTER.
 PA (CNRS) CNRS CENT NAT RECH SCI.
 XX
 PI Duchaud E, Taourit S, Glaser P, Frangeul L, Kunet F, Danchin A;
 PI Buchrieser C;
 DR WPI; 2003-148459/14.
 XX
 PT Genomic sequence of Photorhabdus luminescens and encoded polypeptides,
 PT useful e.g. as therapeutic antimicrobials and agricultural pesticides.
 XX
 PS Claim 2; SEQ ID NO 2163; 1205bp; French.
 XX
 CC The invention relates to the isolation of genes and their encoded
 CC proteins from Photorhabdus luminescens. The isolated sequences are
 CC sources of probes and primers for detecting the genome of P. luminescens
 CC and related species; to study polymorphisms; for gene analysis and for
 CC detection/amplification of the genes. Antibodies (Ab) raised against the
 CC polypeptides encoded by the genes are used for detection/identification
 CC of P. luminescens, e.g. in foods. The genes, proteins, Ab and cells that
 CC carry a gene-containing vector are used to select compounds that
 CC modulate, regulate, induce or inhibit expression of the genes in plants,
 CC animals or microorganisms other than P. luminescens and are able to alter
 CC response or sensitivity to toxins and antibiotics produced by P.
 CC luminescens. Cells transformed to express the genes are useful for
 CC recombinant production of the proteins, particularly toxins and
 CC antibacterials useful as insecticides, bactericides and fungicides. The
 CC genes, proteins, vectors containing the genes and Ab are also useful
 CC therapeutically (to treat microbial infection by bacteria or fungi that
 CC are sensitive to P. luminescens-encoded toxins or antibiotics) and as
 CC biopesticides. Other uses of the genes and the proteins are as virulence
 CC factors and for identifying targets of human diseases for which P.
 CC luminescens is a model (particularly plague and whooping cough). This
 CC sequence represents one of the isolated P. luminescens proteins
 CC
 XX
 SQ Sequence 261 AA;

Query Match 15.6%; Score 197.5; DB 6; Length 261;
 Best Local Similarity 26.3%; Pred. No. 6.2e-14;
 Matches 64; Conservative 44; Mismatches 110; Indels 25; Gaps 7;

QY 10 SLEVSPPKALPPDDKGRYRIAGLKTSLAAGTPIGTLGLVIOGYEMWAPL 69
 DB 8 SMPVSDTSSPLTS-----KKSPFTGIIIDSLPIVGYIPVAFGLNAVRLKGFPMWEIF 62
 QY FSGFLPAGSTEMVIALVVGAAFLGALITLLVFRHVFYAFSPFLVKNPAPRFSYV 129
 DB 70 FSGFLPAGSTEMVIALVVGAAFLGALITLLVFRHVFYAFSPFLVKNPAPRFSYV 129
 DB 63 FSCITVAGASQVITRLLSAGTSLWISALTITMADVRIILVGPL-RRRIDKXLTKEKTV 121
 QY 130 ---PALIDAVAVTAAAR---PAGSVMRLISMQIAFHSYV---VFGLTGVAIAELIP 178
 DB 122 IWAFLTDEVFAAATRAKLIKHRWSSEMMVAIALCSWLWAGATPAAGFLNGYLSYF 181
 QY 179 PEIKGLFPLGLFTLTLDSCRTKKQ-----IPSLLAGSFTIALVPIG--QALFA 230
 DB 188 AIEAMITWPLPFLSFLASCRKNSYVAATLALGALLGITFFSIPAILAGIVGCI 241
 QY 231 ALV 233
 DB 242 ALL 244

RESULT 6
 ABP57756

ID ABP57756 standard; protein; 245 AA.
 AC ABP57756;
 XX
 DT 29-JAN-2003 (first entry)
 XX
 DE E. coli L-amino acid producing protein #1.
 XX
 KM L-amino acid; E. coli.
 XX
 OS Escherichia coli.
 XX
 PN EP1239041-A2.
 XX
 PD 11-SEP-2002.
 XX
 PF 13-FEB-2002; 2002EP-00003335.
 XX
 PR 13-FEB-2001; 2001RU-00103865.
 XX
 PR 26-FEB-2001; 2001RU-00104998.
 XX
 PR 26-FEB-2001; 2001RU-00104999.
 XX
 PR 28-JUN-2001; 2001RU-00117632.
 XX
 PR 28-JUN-2001; 2001RU-00117633.
 XX
 PA (AJIN) AJINOMOTO CO INC.
 XX
 PI Tabolina EA, Rybak KV, Khourges EM, Voroshilova EB, Gussyatiner MM;
 XX
 DR WPI; 2002-69856/76.
 XX
 DR N-P8DB; ABV75656.
 XX
 PT Novel L-amino acid producing Escherichia bacterium, is modified to
 PT enhance L-amino acid production by enhancing the activities of protein
 PT capable of making bacterium to have enhanced resistance to L-amino acids.
 XX
 XX disclosure; Page 23; 33pp; English.
 XX
 CC The invention relates to a novel L-amino acid producing bacterium of
 CC genus Escherichia, modified to enhance L-amino acid production by
 CC enhancing the activities of a protein. The novel bacterium is useful for
 CC producing L-amino acid e.g. L-Thr, L-Val, L-Pro, L-Leu, L-Met and L-Arg,
 CC by cultivating the bacterium in a culture medium and collecting L-amino
 CC acid to be produced and accumulated from the culture medium. The present
 CC sequence represents a protein of the invention which causes increased L-
 CC amino acid production in E. coli
 CC
 XX
 SQ Sequence 245 AA;

Query Match 15.5%; Score 197; DB 5; Length 245;
 Best Local Similarity 29.1%; Pred. No. 6.5e-14;
 Matches 68; Conservative 41; Mismatches 95; Indels 30; Gaps 9;

QY 34 QGKTLAAGLGMVPIGIFGLLVIOGYEMWAPLPSGLIFAGSTEMVIALVVGAAFL 93
 DB 17 EGKGDLPVISTYIPAFAGFNATRLGSPLESYFSCITVAGASQVITRLLSAGSL 76
 QY 94 GATLTLLVFRHVFYAFSPFLHV---KNPIARFYSVFPALIDEAVAVTAAAR---PA 145
 DB 77 WIALVVMADVHVLYGSLRSRIQRLQSKTALW-AFGITDEVFAAATRAKLVRRNR 134
 QY 146 GMSAMRLISMQIAFHSYVFP---GGTLGVAIAELIPFRIKGLFPLGLFTLTLDSC 201
 DB 135 RMEENWMIIGIAFSWSWFGVITGAFSGSLGQIPAVEAALGFMLPFLSFLASIQ 194
 QY 202 TKKQIPSLLAGSFTIALV-VIPQALPA---ALV--IFLGLT--IRYFPLG 247
 DB 195 RKQ-----SLCVTRALVGLAVGLTFSIPAILAGIVCGCLTALIQAFWQ 240

RESULT 7
 ADA35238
 ID ADA35238 standard; protein; 266 AA.

AC ADA35238;
 XX
 DT 20-NOV-2003 (first entry)
 XX
 DE Acinetobacter baumannii protein #2399.
 XX
 DE Acinetobacter baumannii; bacterial disease; antibacterial; vaccine;
 XX
 KW Acinetobacter baumannii; bacterial disease; antibacterial; vaccine;
 XX
 KW plant biocontrol agent.
 XX
 OS Acinetobacter baumannii.
 XX
 PN US6562958-B1.
 XX
 PD 13-MAY-2003.
 XX
 PF 04-JUN-1999; 99US-00328352.
 XX
 PR 09-JUN-1998; 98US-008701P.
 XX
 PA (GENO-) GENOME THERAPEUTICS CORP.
 XX
 PI Breton G, Bush D;
 XX
 DR WPI, 2003-576092/54.
 XX
 DR N-PSDB; ADA31112.
 XX
 PT New Acinetobacter baumannii proteins and nucleic acids, useful as reagents
 PT for diagnosing a bacterial disease, as components of antibacterial
 PT vaccines, as targets for antibacterial drugs, or as biocontrol agents for
 PT plants.
 PS Example; SEQ ID NO 6525; 328pp; English.
 XX
 PS The invention relates to isolated Acinetobacter baumannii nucleic acids.
 CC The A. baumannii nucleic acids and polypeptides are useful as reagents
 CC for diagnosing a bacterial disease, as components of antibacterial
 CC vaccines, as targets for antibacterial drugs, to detect the presence of
 CC A. baumannii and other Acinetobacter species in a sample, in screening
 CC compounds for the ability to interfere with the A. baumannii life cycle
 CC or to inhibit A. baumannii infection, and as biocontrol agents for
 CC plants. The present sequence represents the amino acid sequence of an A.
 CC baumannii protein.
 CC
 SO Sequence 266 AA;
 SQ
 Query Match 15.0%; Score 190; DB 6; Length 266;
 Best Local Similarity 26.6%; Pred. No. 4.6e-13;
 Matches 62; Conservative 43; Mismatches 98; Indels 30; Gaps 7;
 QY 34 QGLKTSIAAGWYPIGIAFGILVIOGYEWMAAPLPSGLIFAGSTEMVIALVVGAAPL 93
 DB 36 QGANDSQAIYVTVTFVGFARFVSASQGFTEWMAFPLSCSWYAGASQFLVVALIASSSI 95
 QY 94 GAIALTLLVNFHFVFA-----FSPELVAVKPIARFYSVPLIDENAA-----VTAAR 143
 DB 96 WLTAITVIALDIRHVLGPALYNLIPKLNKKTAV-----WAWGLTBVVASGNTQISOR 151
 QY 144 PAGSAMRLISMOIAFHSYV---VFGGLTGVAIAELIPFEIKGFALCSLFEVTLTLD 199
 DB 152 ROOMSESMGLISLFSWMSWASGSLIGLFPADQVAHLPKFLQALDLPRLPLFSFLAA 211
 QY 200 CRITKQIPLLLAGISFTIALVTPGQALFPAALIFLGLT-----IRYFFL 246
 DB 212 FERKHS---LVVAVSLVSAALACYMINTLSAA--IFGIISGLIAGLGFXYVL 258
 RESULT 8
 ADA33199 standard; protein; 263 AA.
 ID ADA33199
 AC ADA33199;
 XX
 DT 20-NOV-2003 (first entry)

XX
 DE Acinetobacter baumannii protein #360.
 XX
 DE Acinetobacter baumannii; bacterial disease; antibacterial; vaccine;
 XX
 KW Acinetobacter baumannii; bacterial disease; antibacterial; vaccine;
 XX
 KW plant biocontrol agent.
 XX
 OS Acinetobacter baumannii.
 XX
 PN US6562958-B1.
 XX
 PD 13-MAY-2003.
 XX
 PF 04-JUN-1999; 99US-00328352.
 XX
 PR 09-JUN-1998; 98US-008701P.
 XX
 PA (GENO-) GENOME THERAPEUTICS CORP.
 XX
 PI Breton G, Bush D;
 XX
 DR WPI, 2003-576092/54.
 XX
 DR N-PSDB; ADA29073.
 XX
 PT New Acinetobacter baumannii proteins and nucleic acids, useful as reagents
 PT for diagnosing a bacterial disease, as components of antibacterial
 PT vaccines, as targets for antibacterial drugs, or as biocontrol agents for
 PT plants.
 PS Example; SEQ ID NO 4486; 328pp; English.
 XX
 PS The invention relates to isolated Acinetobacter baumannii nucleic acids.
 CC The A. baumannii nucleic acids and polypeptides are useful as reagents
 CC for diagnosing a bacterial disease, as components of antibacterial
 CC vaccines, as targets for antibacterial drugs, to detect the presence of
 CC A. baumannii and other Acinetobacter species in a sample, in screening
 CC compounds for the ability to interfere with the A. baumannii life cycle
 CC or to inhibit A. baumannii infection, and as biocontrol agents for
 CC plants. The present sequence represents the amino acid sequence of an A.
 CC baumannii protein.
 CC
 SO Sequence 263 AA;
 SQ
 Query Match 14.9%; Score 189.5; DB 6; Length 263;
 Best Local Similarity 24.0%; Pred. No. 5.2e-13;
 Matches 49; Conservative 54; Mismatches 96; Indels 5; Gaps 5;
 QY 40 LAAAGWYPIGIAFGILVIOGYEWMAAPLPSGLIFAGSTEMVIALVVGAAPLGAIALT 99
 DB 53 LPLSIVIPMALIAGSMVHGLSFKXALMSGIVFAQAQVLSMVEGASLITIVT 112
 QY 100 TLLVNFHFVFAFSPELVAVKPIARFYSV-PALIDENAYTAARPAQMSAMRLISMOIA 158
 DB 113 IFFLTAQHFTVIALTRNDISILPLSKRLTGLLDELFAVSVPEKKGPOI-LGAGIIC 171
 QY 159 FHSWVFGGLTGVAIAELIPFEIK-GLFEALCSLFEVTLTLDSCRTKQIPLLLAGLS-F 216
 DB 172 FIDFWVPSLVGILATLALPILNLNHLDPSTIAIVAMIVPCKCKPWWAGILMTCVSGF 231
 QY 217 TIALVTPGQALFPAAL-IFLGLL 239
 DB 232 VLKFEHIEGAILISGLHGFIAVI 255
 RESULT 9
 ADC95293 standard; protein; 255 AA.
 ID ADC95293
 AC ADC95293;
 XX
 DT 01-JAN-2004 (first entry)
 XX
 DE E. faecium protein sequence SEQ ID 4920.
 XX

KW Vaccine; urinary tract infection; bacteraemia; endocarditis; wound;
 KW abdominal-pelvic infection.
 OS Enterococcus faecium.
 XX US6583275-B1.
 XX 24-JUN-2003.
 XX 30-JUN-1998; 98US-00107532.
 XX 02-JUN-1997; 97US-0051571P.
 XX 14-MAY-1998; 98US-0085598P.
 XX (GENO-) GENOME THERAPEUTICS CORP.
 XX Doucetbe-Stamm LA, Bush D;
 XX MPI; 2003-799836/75.
 DR N-PSDB; ADO91639.
 XX New isolated nucleic acid derived from Enterococcus faecium encoding an
 PT Enterococcus faecium polypeptide useful for detection, prevention and
 PT treatment of a pathological condition resulting from a bacterial
 PT infection.
 XX Example 1; SEQ ID NO 4920; 243pp; English.
 XX The invention relates to an isolated nucleic acid derived from
 CC Enterococcus faecium encoding an Enterococcus faecium polypeptide having
 CC one of 10 fully defined sequences given in the (or comprising 40
 CC sequential nucleotides chosen from any of the nucleic acids, its
 CC complement or sequences hybridizing to it). Also included are a
 CC recombinant vector comprising the nucleic acid operably linked to
 CC transcription regulatory element, a cell comprising the vector and a
 CC single-stranded probe comprising the nucleic acid. The nucleic acids are
 CC chosen from 3654 disclosed sequences encoding 3654 disclosed proteins.
 CC The nucleic acids are useful for diagnosing pathological conditions
 CC resulting from E. faecium bacterial infection (e.g. urinary tract
 CC infection, bacteraemia, endocarditis, wounds and abdominal-pelvic
 CC infection) and for screening drugs such as agonists and antagonists. The
 CC nucleic acid is useful for recombinant production of Candida albicans -
 CC derived peptides or antisense polypeptides. Pharmaceutical compositions
 CC and vaccines containing the nucleic acid are useful for preventing or
 CC treating Enterococcus faecium infections. The present sequence represents
 CC one if the disclosed E. faecium proteins.
 XX Sequence 255 AA;
 SQ
 Query Match 14.6%; Score 185.5; DB 7; Length 255;
 Best Local Similarity 25.0%; Pred. No. 1.4e-11;
 Matches 63; Conservative 51; Mismatches 105; Indels 33; Gaps 10;
 QY 16 SKAALEPPDKGRRREINQGLKTSIAGLGNTPYFGIAGFLVLYQYGVMAAPLPSGLIF 75
 DB 15 AKGVALLDERK---LDITAKIKDTLPVFGYTGIGLAFVIGKRAGHPVVTMLSLVY 70
 QY 76 AGSTEMLVIALVGAAPGATATLTLLVNFHVFYAFSPFLHVVYKQPIARFVSVALI-D 134
 DB 71 AGSAQFIVSMASHPSLISIVSFSTLVNRMIMSWTIAFYFKNLLQNLIGTLMD 130
 QY 135 EAYAVTARPAQWMSMRISQIAFH-----SY--WYFGGTLGYAIELPFIK-G 183
 DB 131 ESFAL-----GNNALNTTGGKLNFRMMNTNMISYLTWSSSLVGLLGNFTIDPKFG 184
 QY 184 LEFALCSLFTVLT---TLDSCTKXQPIPSLLAGLSFT---IALVPIGQALFALLIF- 235
 DB 185 LDFALVAMFGLLYQVSDRNTSKRL-QLITLIGTLILVVGILFIPSNLIVVVTLLG 243
 QY 236 --IGLLTRPYF 245
 DB 244 CGLGVWTKHAF 255

RESULT 10
 ID ABU00481
 XX ABU00481 standard; protein; 218 AA.
 XX AC ABU00481;
 XX 23-OCT-2003 (revised)
 DT 11-FEB-2003 (first entry)
 XX S. pneumoniae type 4 strain protein from coding region #48.
 XX Bacterial meningitis; pneumonia; sepsis; otitis media; ear infection;
 KW antiinflammatory; antibacterial; immunostimulant; auditory; respiratory;
 KW gene therapy; vaccine.
 XX Streptococcus pneumoniae; type 4 strain.
 OS WO200277021-A2.
 XX 03-OCT-2002.
 XX 27-MAR-2002; 2002WO-1B002163.
 XX 27-MAR-2001; 2001GB-00007658.
 XX (CHIR-) CHIRON SPA.
 PA (GENO-) INST GENOMIC RES.
 XX Maignani V, Tettelin H, Fraser C;
 DR MPI; 2003-040579/03.
 DR N-PSDB; ABX05760.
 XX New proteins and nucleic acid molecules from Streptococcus pneumoniae,
 PT useful as medicaments for treating or preventing a disease or infection
 PT due to streptococcus bacteria, such as pneumonia, sepsis, otitis media or
 PT ear infection.
 XX Claim 1; SEQ ID NO 96; 56pp; English.
 XX The invention relates to a protein comprising or having at least 50%
 CC identity to any of the 2469 amino acid sequences, identified in the
 CC specification (available on a computer readable format), or its fragment,
 CC expressed from 2469 identified DNA coding regions from the
 CC Streptococcus pneumoniae type 4 strain genomic sequence appearing as
 CC ABS6444. Also included are an antibody which binds one of the proteins,
 CC treating a patient by administering the protein, DNA or antibody (in a
 CC composition), a kit comprising first and second primers, which are the
 CC nucleic acid cited above or fragments between nucleotides 8-100 of a
 CC sequence not defined in the specification, for amplifying a target
 CC sequence contained within a Streptococcus nucleic acid sequence, where
 CC the first primer is substantially complementary to the target sequence
 CC and the second primer is substantially complementary to the complement of
 CC the target sequence, and where the pairs of the primers having
 CC substantial complementarity define the termini of the target sequence to
 CC be amplified, assay comprising contacting a test compound binds to the
 CC protein, and determining whether the test compound binds to the protein
 CC and a Streptococcus pneumoniae bacterium, where one or more genes
 CC encoding the proteins has been rendered inactive. The proteins, nucleic
 CC acid molecules, antibody and compositions are useful as medicaments for
 CC treating or preventing a disease or infection due to streptococcus
 CC bacteria, particularly S. pneumoniae, such as pneumonia, sepsis, otitis
 CC media or ear infection. They are also useful in developing vaccines,
 CC immunostimulant proteins. The methods are useful for identifying
 CC immunodominant proteins. The present sequence is one of the 2469 proteins
 CC expressed by the identified coding regions from the genomic sequence.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences. (Updated on 23-OCT-2003 to
 CC standardise OS field)
 XX Sequence 218 AA;
 SQ

Query March 14.0%; Score 177.5; DB 6; Length 218;
Best Local Similarity 28.6%; Pred. No. 9, 7e-12;
Matches 63; Conservative 35; Mismatches 82; Indels 39; Gaps 11;

QY 40 LAAGLGNYPGIGAFGLIVIQYGYEMWAPLPFSGL---ITAGSTEMVIALVWGAPLGA 95
DQ 1 MPTALGVASIGLGGIIGAPY-----VTPVEMGIMSLFVYAGSOFMTALIVQAAVAA 55
QY 96 IATLTLLVNRHAFVAFSPFLHYVK--NPVAFPSV---PALDEAYAVAAPAG---- 146
DQ 56 IATVTELLNR---LFLSLHASTYFRHFSIMINIGMSILTDITYGVLNGELAHFDKV 111
QY 147 ---WSAMRLISMQIAFHSYVWFGGLNGVAIAELIPF-EIKGLEBALCSLFVTLTDSGR- 201
DQ 112 NPMNMHGNHNNINSYVA---WFVGTVGTFALGGLLENPEIFGLDPAIVGMFIGIPASOFQI 167
QY 202 TKKQIP-----SLLAGLSFTALVYIPQ--ALFAALL 233
DQ 168 MQRIPVNRLLIILAVAVSFFLLITVMSGLAVFATLL 207

RESULT 12
ABBS4959
ID ABBS4959 standard; protein: 235 AA.

XX ABB54959;
AC
XX 29-AUG-2003 (revised)
DT
XX 16-MAY-2002 (first entry)
DT
XX Lactococcus lactis protein yqfD.
DE
XX Biosynthesis; biodegradation; lactic bacterium; yogurt; cheese.
KM
XX Lactococcus lactis; IL1403.
OS
XX FR2807446-A1.
FN
XX 12-OCT-2001.
PD
XX 11-APR-2000; 2000FR-00004630.
PF
XX 11-APR-2000; 2000FR-00004630.
PR
XX (INRG) INRA INST NAT RECH AGRONOMIQUE.
PA
XX Bolotine A, Sorokine A, Renault P, Ehrlich SD;
PI
XX WPI; 2002-043418/06.
DR
XX
XX New nucleotide sequence useful in the identification of Lactococcus
XX lactic and related species.
XX
XX Claim 6; SEQ ID NO 1661; 2504bp; French.

XX The present invention is related to a Lactococcus lactis nucleotide
XX sequence (ABBS4959) and related proteins (ABBS3300-ABBS5621). The nucleic
XX acid sequence is useful in the detection and/or amplification of nucleic
XX acid sequence, particularly to identify Lactococcus lactis or related
XX species. The proteins of the invention are useful for the biosynthesis or
XX biodegradation of a composition of interest. The invention helps research
XX in lactic bacteria, particularly useful in the production of yogurt and
XX cheese. Note: The sequence data for this patent is based on equivalent
XX patent WO2001/7334 (published 18-OCT-2001) which is available in
XX ftp://wipo.int/pub/published_pct_sequences. (updated on 29-AUG-2003 to
XX standardise OS field)
XX
XX Sequence 235 AA:
SQ

Query Match 13.8%; Score 174.5; DB 5; Length 235;
Best Local Similarity 25.5%; Pred. No. 2, 4e-11;

CC sequence and proteins encoded by it are useful in pharmaceutical and

PS Claim 1; Page 4153; 4525pp; English

CC The invention relates to a protein (ABP2413-ABP30895) from group B
CC streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS
CC (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in

XX

Search completed: April 16, 2004, 06:53:11
Job time : 73.6156 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Comphen Ltd.

OM protein - protein search, using sw model

Run on: April 16, 2004, 06:51:17 / Search time 23.716 Seconds
(without alignments)

545.110 Million cell updates/sec

Title: US-10-608-504-3

Perfect score: 1269
Sequence: 1 MKOTGHSLSVSPSKAL.....LIIFGLTIRYFLGKAAK 251

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database: Issued Patents AA:
1: /cgn2_6/ptodata/2/1aa/5A_COMB.pep.*
2: /cgn2_6/ptodata/2/1aa/5B_COMB.pep.*
3: /cgn2_6/ptodata/2/1aa/6A_COMB.pep.*
4: /cgn2_6/ptodata/2/1aa/6B_COMB.pep.*
5: /cgn2_6/ptodata/2/1aa/PTUS_COMB.pep.*
6: /cgn2_6/ptodata/2/1aa/backfile1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1269	100.0	251	US-09-471-803A-3	Sequence 3, Appl
2	221.5	17.5	247	US-09-489-039A-8478	Sequence 8778, Ap
3	190	15.0	266	US-09-328-352-6525	Sequence 6525, Ap
4	189.5	14.9	263	US-09-328-352-4486	Sequence 4486, Ap
5	185.5	14.6	255	US-09-107-532A-4920	Sequence 4920, Ap
6	181	14.3	251	US-09-489-039A-10187	Sequence 10187, A
7	175	13.8	345	US-09-252-991A-27947	Sequence 27947, A
8	171.5	13.5	288	US-09-489-039A-12218	Sequence 12218, A
9	161.5	12.5	242	US-09-134-001C-5154	Sequence 5154, Ap
10	141.5	11.2	247	US-09-328-352-6024	Sequence 6024, Ap
11	137	10.8	206	US-09-134-000C-6789	Sequence 6789, Ap
12	134.5	10.6	230	US-09-543-681A-4173	Sequence 4173, Ap
13	113.5	8.9	484	US-09-252-991A-32299	Sequence 32299, A
14	106	8.4	270	US-09-252-991A-18185	Sequence 18185, A
15	104	8.2	555	US-09-252-991A-31521	Sequence 31521, A
16	102.5	8.1	516	US-09-252-991A-21329	Sequence 21329, A
17	100	7.9	506	US-09-540-236-2360	Sequence 2360, Ap
18	98.5	7.8	642	US-08-706-844-32	Sequence 32, Appl
19	98.5	7.8	643	US-08-706-844-39	Sequence 39, Appl
20	98.5	7.8	643	US-08-599-654-39	Sequence 39, Appl
21	98.5	7.8	643	US-08-944-868A-39	Sequence 39, Appl
22	98.5	7.8	643	US-08-944-423A-39	Sequence 39, Appl
23	98.5	7.8	643	US-08-944-486-39	Sequence 39, Appl
24	97	7.6	345	US-09-543-681A-7546	Sequence 7546, Ap
25	96.5	7.6	417	US-08-553-888A-3	Sequence 3, Appl
26	96.5	7.6	490	US-09-252-991A-27210	Sequence 27210, A
27	95.5	7.5	452	US-09-252-991A-19116	Sequence 19116, A

28	95	7.5	1334	2	US-08-996-545-2	Sequence 2, Appl
29	95	7.5	1334	3	US-09-328-320-2	Sequence 2, Appl
30	94	7.4	421	4	US-09-489-039A-7699	Sequence 7699, Ap
31	94	7.4	434	4	US-09-543-681A-7154	Sequence 7154, Ap
32	93.5	7.4	486	4	US-09-291-922-10	Sequence 10, Appl
33	93.5	7.4	591	4	US-09-540-236-2490	Sequence 2490, Ap
34	92.5	7.3	326	4	US-09-107-532A-3665	Sequence 3665, Ap
35	91.5	7.2	274	4	US-09-489-039A-14140	Sequence 14140, A
36	91	7.2	472	4	US-09-489-039A-10315	Sequence 10315, A
37	90.5	7.1	532	4	US-09-252-991A-22421	Sequence 22421, A
38	90	7.1	490	4	US-09-252-991A-30355	Sequence 30355, A
39	88.5	7.0	484	4	US-09-489-039A-11054	Sequence 11054, A
40	88	6.9	2864	4	US-08-469-260A-394	Sequence 394, App
41	88	6.9	2864	4	US-08-467-344A-394	Sequence 394, App
42	88	6.9	2864	4	US-08-467-344A-394	Sequence 394, App
43	87.5	6.9	492	4	US-09-134-001C-4847	Sequence 4847, Ap
44	87.5	6.9	516	4	US-09-489-039A-12893	Sequence 12893, A
45	87.5	6.9	623	4	US-09-252-991A-19867	Sequence 19867, A

ALIGNMENTS

RESULT 1
US-09-471-803A-3
Sequence 3, Application US/09471803A

Patent No. 6613545

GENERAL INFORMATION:
APPLICANT: KENNERKNECHT, NICOLE

APPLICANT: SAHM, HERMANN

APPLICANT: EGGELING, LOTAR

APPLICANT: PFEFFELBE, WALTER

TITLE OF INVENTION: NUCLEOTIDE SEQUENCE CODING FOR THE EXPORT OF
BRANCHED CHAIN AMINO ACIDS. PROCESS FOR THE

TITLE OF INVENTION: ISOLATION THEREOF AND USE THEREOF

FILE REFERENCE: 21123/265496/MAS

CURRENT FILING DATE: 1999-12-23

PRIOR APPLICATION NUMBER: DE 199 51 708.8

PRIOR FILING DATE: 1999-10-27

NUMBER OF SEQ ID NOS: 12

SOFTWARE: Patentin Ver. 2.1

SEQ ID NO 3

LENGTH: 251

TYPE: PRT

ORGANISM: Corynebacterium glutamicum

FEATURE:
OTHER INFORMATION: ATCC14752

US-09-471-803A-3

Query Match

Best Local Similarity

Matches 251, Conservative 0, Mismatches 0, Indels 0, Gaps 0;

QY	1	MKOTGHSLSVSPSKALBPDDGGRYRYEAGLKTSLAAGLGMYPFGIAFGLVTOY	60
DB	1	MKOTGHSLSVSPSKALBPDDGGRYRYEAGLKTSLAAGLGMYPFGIAFGLVTOY	60
QY	61	GYYWMAAPLFSGLIFAGSTEMLVIALVGAAPLGAIALTTLLVNRHVFYAFSPFLAYVK	120
DB	61	GYYWMAAPLFSGLIFAGSTEMLVIALVGAAPLGAIALTTLLVNRHVFYAFSPFLAYVK	120
QY	121	NPTARYSPALIDEXYANTARPPGMSMRLISQIAFHSYVWGGLTGVAIAELIPE	180
DB	121	NPTARYSPALIDEXYANTARPPGMSMRLISQIAFHSYVWGGLTGVAIAELIPE	180
QY	181	IKGLEFALCSLFTVTLTDSCTKQIPSLLAGLSTFTALVVIPOQALFALLIFGLLT	240
DB	181	IKGLEFALCSLFTVTLTDSCTKQIPSLLAGLSTFTALVVIPOQALFALLIFGLLT	240
QY	241	IRYFPLGKAAK 251	
DB	241	IRYFPLGKAAK 251	

RESULT 2

US-09-489-039A-8478
 ; Sequence 8478, Application US/09489039A
 ; Patent No. 6610836
 ; GENERAL INFORMATION:
 ; APPLICANT: Gary Breton et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
 ; FILE REFERENCE: 2709.2004001
 ; CURRENT APPLICATION NUMBER: US/09/489,039A
 ; PRIOR FILING DATE: 2000-01-27
 ; PRIOR FILING DATE: 1999-01-29
 ; NUMBER OF SEQ ID NOS: 14342
 ; SEQ ID NO 8478
 ; LENGTH: 247
 ; TYPE: PR
 ; ORGANISM: Klebsiella pneumoniae
 US-09-489-039A-8478

Query Match 17.5%; Score 221.5; DB 4; Length 247;
 Best Local Similarity 28.0%; Pred. No. 2.4e-17;
 Matches 61; Conservative 41; Mismatches 101; Indels 15; Gaps 6;

QY 32 IAGGKTSLAGLGMVPIGAFGLVIOGYEWMAAPLFGSLIFAGSTEMLVIALVVGAA 91
 DB 7 IFRGASAIMPLCIDPSPFIVGALVSAGSWQSTAMSAIVAGSOMLALNMLKTGA 66
 QY 92 PLGIALTLTLVNRHVFYAFSPFLHYVKNPIAREFS---VFALIDEAVATTAAPAGW 147
 DB 67 TLGVITFTLTINRHVYASISGIVRE---ASFCKCFMSYALTDVAVATTKMEGN 123
 QY 148 SAMRLI---SMQIAFHSYVWFGGLTGVAAIAELIP-PEIKGLEFALCSLFTLTLDSCRTK 203
 DB 124 KKEKYLFGSAMTFWMAIWLADFLGALVGSFHIKYGGLDFAMVAFAIIVVPOIKSQ 183
 QY 204 KQISLILAGLSTITALVIVP---GQALFALLFLDL 238
 DB 184 ACTVAAVVAVS-GTLVLVLPYSLGIIVASVGLVAGL 220

RESULT 3

US-09-328-352-6525
 ; Sequence 6525, Application US/09328352
 ; Patent No. 6562958
 ; GENERAL INFORMATION:
 ; APPLICANT: Gary L. Breton et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
 ; FILE REFERENCE: GTC99-03PA
 ; CURRENT APPLICATION NUMBER: US/09/328,352
 ; PRIOR FILING DATE: 1999-06-04
 ; NUMBER OF SEQ ID NOS: 8252
 ; SEQ ID NO 6525
 ; LENGTH: 266
 ; TYPE: PR
 ; ORGANISM: Acinetobacter baumannii
 US-09-328-352-6525

Query Match 15.0%; Score 190; DB 4; Length 266;
 Best Local Similarity 26.6%; Pred. No. 1.1e-13;
 Matches 62; Conservative 43; Mismatches 98; Indels 30; Gaps 7;

QY 34 GGLKTSIAGLGMVPIGAFGLVIOGYEWMAAPLFGSLIFAGSTEMLVIALVVGAA 93
 DB 36 QAKSQAIVLTVLSPFVGSASQFETPEAFSLSCSMYAGASQFLVALLASGSI 95
 QY 94 GAIALTTLLVNRHVFYAFSPFLHYVKNPIAREFSVFALIDEAIA---VTAR 143
 DB 96 WTLATLVIALDIRHVLVYGPALYNLPTKMLKTKAV---WAMGLTDEVFASGMQLSOR 151

QY 144 PAGSAMLLISMQIAFHSYV---VFGGLTGVAAIAELIPPEIKLEPALSFLVTLDS 199
 DB 152 RQSWSEWMLGSLFSWMSWASGSLGLFPADGVVHLPKFLQALDPLALFLSLIAA 211
 QY 200 CRTKQIPSLILAGLSTITALVIVPQALFPALLIFGLT-----IRYFEL 246
 DB 212 FERKHS---LVAVSLVSVSLACTWINLSAA--IFIGILSGIAGLFXKYVL 258

RESULT 4

US-09-328-352-4486
 ; Sequence 4486, Application US/09328352
 ; Patent No. 6562958
 ; GENERAL INFORMATION:
 ; APPLICANT: Gary L. Breton et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
 ; FILE REFERENCE: GTC99-03PA
 ; CURRENT APPLICATION NUMBER: US/09/328,352
 ; PRIOR FILING DATE: 1999-06-04
 ; NUMBER OF SEQ ID NOS: 8252
 ; SEQ ID NO 4486
 ; LENGTH: 263
 ; TYPE: PR
 ; ORGANISM: Acinetobacter baumannii
 US-09-328-352-4486

Query Match 14.9%; Score 189.5; DB 4; Length 263;
 Best Local Similarity 24.0%; Pred. No. 1.3e-13;
 Matches 49; Conservative 54; Mismatches 96; Indels 5; Gaps 5;

QY 40 IAGGKMPYIGAFGLVIOGYEWMAAPLFGSLIFAGSTEMLVIALVVGAAPIGAILT 99
 DB 53 LPLSIVIPMALIASMAVHGLSFKYKALNSGIVFAGAQLVLSVMWEGASLITIVT 112
 QY 100 TLVNRHVFYAFSPFLHYVKNPIAREFSV-FALIDEAVATTAAPAGMSAMRLISMQIA 158
 DB 113 IFFLTQHFITVLTNRNDISILPSKRLTGLTDELFAVSVPEKHPQY-LFGAGLC 171
 QY 159 FHSWVFGGLTGVAAIAELIPPEIK-GLFPALCSLFTLTLDSCRTKQIPSLILAGLS-F 216
 DB 172 FTLFWVSVSLVILATLALPILNTHDPSITAIIVAVIWPCKCKPVMAGILMTCVSGF 231
 QY 217 TIALVVPQALFALL-IFGLL 239
 DB 232 VLKFFHIEGALLISGLMGFIAVI 255

RESULT 5

US-09-107-532A-4920
 ; Sequence 4920, Application US/09107532A
 ; Patent No. 6583275
 ; GENERAL INFORMATION:
 ; APPLICANT: Lynn A Doucette-Stamm and David Bush
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
 ; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
 ; NUMBER OF SEQUENCES: 7310
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSER: GENOME THERAPEUTICS CORPORATION
 ; STREET: 100 Beaver Street
 ; CITY: Waltham
 ; STATE: Massachusetts
 ; COUNTRY: USA
 ; ZIP: 02354
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: CD/ROM ISO9660
 ; COMPUTER: PC
 ; OPERATING SYSTEM: <Unknown>
 ; SOFTWARE: ASCII
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/107,532A
 ; FILING DATE: 30-Jun-1998
 ; PRIOR APPLICATION DATA:


```

APPLICATION NUMBER: 60/085,598
FILING DATE: 14 May 1998
APPLICATION NUMBER: 60/051571
FILING DATE: July 2, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Arinello, Pamela Deneke
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-012
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781)893-5007
TELEFAX: (781)893-8277
INFORMATION FOR SEQ ID NO: 4920:
SEQUENCE CHARACTERISTICS:
LENGTH: 255 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHEetical: YES
ORIGINAL SOURCE:
ORGANISM: Enterococcus faecium
FEATURE:
NAME/KEY: misc feature
LOCATION: (B) LOCATION 1...255
SEQUENCE DESCRIPTION: SEQ ID NO: 4920:
US-09-107-532A-4920

Query Match
Best Local Similarity 14.6%; Score 185.5; DB 4; Length 255;
Matches 63; Conservative 51; Mismatches 105; Indels 33; Gaps 10;

16 SKAALEPDDKGRYRVEIAQGLKTSIAGLGMYPPIGAFGLVLYQGYEMWAAPLFSGILF 75
15 AKGVLLDKR---LDIKTAIKDTLPTVGYIGIGAFVIGKAAGFHVLVTLMSLIVY 70
76 AGSTEMVLAVVGAPFLGALITLLNFRHVFYAFSPFLVKNPIAFSVALLI-D 134
71 AGSAQFTVSMASHSPLSIVSFITLVNSRMLMSMTAFYFKKRLQNLITGLTLD 130
135 EAVAVVAAPAGWSARLISMOJAFH-----SY--VWFGGLGVAIAELIPEIK-G 183
131 ESPAL-----GWNKNTNYGQKLNFRMNTAMWISYLVWGSGLVALLGNFTDPKRG 184
184 LEFALCSLFTVLT---TLDSCKTKQIPLSLLAGLSFT---IALVYIPQALFALLIF- 235
185 LDFAIYAFMFLIGLIVLVISDRNTSKRL-QLIILGLTLLVYGLIFIPSNLVIVVTLIG 243
236 --LGLLTIRYFF 245
244 CGLGVWIKHAF 255
DB

RESULT 6
US-09-489-039A-10187
Sequence 10187, Application US/09489039A
Patent No. 6610835
GENERAL INFORMATION:
APPLICANT: Gary Breton et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
FILE REFERENCE: 2709.2004001
CURRENT APPLICATION NUMBER: US/09/489,039A
CURRENT FILING DATE: 2000-01-27
PRIOR APPLICATION NUMBER: US 60/117,747
PRIOR FILING DATE: 1999-01-29
NUMBER OF SEQ ID NOS: 14342
SEQ ID NO 10187
LENGTH: 251
TYPE: PRT
ORGANISM: Klebsiella pneumoniae
US-09-489-039A-10187

Query Match
Best Local Similarity 14.3%; Score 181; DB 4; Length 251;
Matches 68; Conservative 36; Mismatches 95; Indels 28; Gaps 10;

```

```

32 IAQGLKTSIAGLGMYPPIGAFGLVLYQGYEMWAAPLFSGILFASSTMLVIALVGA 91
21 VGGVSDSLPIVSYDPVAFAPGLNATRGFTPELSFSCITIVAGASQVITMLAAGS 80
92 PLGIALTLTNFRHVFYAFSPFLV---VKNPIAFSVFALIDEAAVTAAR---- 143
81 SLWVAALTVMAMDVHVLGPRLRSRLRSALDKKTAAM--AGLTDVFAATATLVVD 138
144 PAGWSA-WRLISQIAFHSY--VWFGGLGV---AIAELIPEIIGLSEFALCSLFTVLT 196
139 NRRWSENMWL---GLAFTSWASWVCGTLGAMSGNGILDYPAVEAALGFMPLFMSEFL 195
197 LDCSTKTKQI-PSLLAGLSFTIALVYIPQALFALLI--IFGLLT 240
196 LASFQROQSLCVTRALAGALGILFSIP-----AIIAGIVCGCLT 237
DB

RESULT 7
US-09-252-991A-27947
Sequence 27947, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 27947
LENGTH: 345
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-27947

Query Match
Best Local Similarity 13.8%; Score 175; DB 4; Length 345;
Matches 52; Conservative 47; Mismatches 104; Indels 12; Gaps 6;

29 RYETIAQGLKTSIAGLGMYPPIGAFGLVLYQGYEMWAAPLFSGILFASSTMLVIALV 88
117 RQAFLRGAVAILPLSLAVAPWGLAGSMATEANLTPAEGQSLAIVPAGAAQVAIGMLK 176
89 GAAPLGAIALTLTNFRHVFYAFSPFLVKNPIAFSV---FALIDEAAVTAAR- 144
177 CGAGFSPSITFTLLITLTSQHLIYGS--LRVLSPLPGRMRIUGFLTLDEFPLAASQHDR 234
145 AGWSAMRLISQIAFHSYWFGGLTVGAIAELIP-PEIKGLEPACSLFTVLTDSCKTK 203
235 RNFNRWYALGVGLFYIYAMNFTLAGILGRSIPGLHGLDFSIATFTALVAPLVR-- 292
204 KQIPSLLAGLSFTIALVYIPQALFALLIFGL 238
293 -NFTLVCAVNTSLCSVILFSHMQ--WSSALVLAQL 324
DB

RESULT 8
US-09-489-039A-12218
Sequence 12218, Application US/09489039A
Patent No. 6610835
GENERAL INFORMATION:
APPLICANT: Gary Breton et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
FILE REFERENCE: 2709.2004001
CURRENT APPLICATION NUMBER: US/09/489,039A
CURRENT FILING DATE: 2000-01-27

```

PRIOR APPLICATION NUMBER: US 60/117,747
 PRIOR FILING DATE: 1999-01-29
 NUMBER OF SEQ ID NOS: 14342
 SEQ ID NO 12218
 LENGTH: 288
 TYPE: PRT
 ORGANISM: Klebsiella pneumoniae
 US-09-489-039A-12218

Query Match 13.5%; Score 171.5; DB 4; Length 288;
 Best Local Similarity 27.5%; Pred. No. 1.8e-11;
 Matches 58; Conservative 41; Mismatches 87; Indels 25; Gaps 10;

QY 40 LAAGLMYPIGIAFGLLVIOGYEWMAPLPSGLIFAGSTEMLVIALVY-GAAPGATL 98
 DB 90 LAAGV---VGMVYGLAAVYGFPLVWPLLSVLAASSEFFIGIVASGSPAA-AA 144
 QY 99 TTVLVFRHVFYAFSPFLHVNKNPIAR---FYSVEALIDEAYAVTAAAP--GMSAMR 151
 DB 145 AGLLVNAHPIFGVT-----VRELVGRRALSFICGHIMDESVFGLSGPTPAQRKAAVW 199
 QY 152 LISMOIAFHSYVWPGGLTGAIAELIP-EIKGLEFALCSLFTLLDSCRTKKQIPSL 210
 DB 200 LCGAGVAL--IMPLGLTGAAGKLPAPETIGLDVFPAILALVPAFKR---TTLI 254
 QY 211 LAGLSFTLIVIPGOAL-FALLIFGLLT 240
 DB 255 RAASGALALAAVFPVPTGLPVLISLFGILS 285

RESULT 9

US-09-134-001C-5154
 Sequence 5154, Application US/09134001C
 Patent No. 6380370
 GENERAL INFORMATION:
 APPLICANT: Lynn Doucette-Stamm et al
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
 FILE REFERENCE: GTC-007
 CURRENT APPLICATION NUMBER: US/09/134,001C
 CURRENT FILING DATE: 1998-08-13
 PRIOR APPLICATION NUMBER: US 60/064,964
 PRIOR FILING DATE: 1997-11-08
 PRIOR APPLICATION NUMBER: US 60/055,779
 PRIOR FILING DATE: 1997-08-14
 NUMBER OF SEQ ID NOS: 5674
 SEQ ID NO 5154
 LENGTH: 242
 TYPE: PRT
 ORGANISM: Staphylococcus epidermidis
 US-09-134-001C-5154

Query Match 12.7%; Score 161.5; DB 4; Length 242;
 Best Local Similarity 24.8%; Pred. No. 2e-10;
 Matches 57; Conservative 40; Mismatches 90; Indels 43; Gaps 8;

QY 34 OGLKSLAAGLGMPIGIAFGLLVIOGYEWMAPLPSGLIFAGSTEMLVIALVGAAPL 93
 DB 20 QGVKECIPITLLGYAGVGISFGVIAASQNFVLEIILLCLLIYAAAGPITITVIAAGTPI 79
 QY 94 GAIATLTLVFRHVFYAFSPFLHVNKNPIARFYSVF-----ALIDEAYAVTAAAP-- 144
 DB 80 SAIVLTTLIVNR-----MFLSWTLAPNPKQYGFNRVGLTITLDETPGV-ATIPYV 132
 QY 145 -----AGSAMRLISMOIAFHSYVWPGGLTGAIAELIP-EIKGLEFALCSLFTLLT 197
 DB 133 KGEKINDRM---LHGNIATVLEFTVSCVIGAFGEYSINPDALGDFATTAFFICI 188
 QY 138 DSCRTKKQ-----IPSLLAGLSFTLIVIPGOALFAALL 233
 DB 189 SQFBEIKKSRLRIYVLIVCVIVMMLLSIIPSVAILI--AAIVALL 236

RESULT 10

US-09-328-352-6024
 Sequence 6024, Application US/09328352
 Patent No. 6562938
 GENERAL INFORMATION:
 APPLICANT: Gary L. Breton et al.
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
 FILE REFERENCE: GTC99-039A
 CURRENT APPLICATION NUMBER: US/09/328,352
 CURRENT FILING DATE: 1999-06-04
 NUMBER OF SEQ ID NOS: 8252
 SEQ ID NO 6024
 LENGTH: 247
 TYPE: PRT
 ORGANISM: Acinetobacter baumannii
 US-09-328-352-6024

Query Match 11.2%; Score 141.5; DB 4; Length 247;
 Best Local Similarity 26.4%; Pred. No. 4.1e-08;
 Matches 52; Conservative 37; Mismatches 95; Indels 13; Gaps 8;

QY 49 IGIAFGLLVIOGYEWMAPLPSGLIFAGSTEMLVIA-LVGAAPLGAIALTTLLVNRH 107
 DB 53 VGMISGLSIAASYGLALMIPCLSLFVLAGTAEFFIFIGFLAVGSPDAA-AAAGLLVNLRH 111
 QY 108 VFYAFSPFLHVNKNPIARFYSVFPALIDEAYAVTAAAP--AGSAMRLISMOIAFHSYV 164
 DB 112 LPFGIAVN-EMIRGFSGQFSGHIMDESVLFMAQDPFETKKAAYWLCGIGILS--WP 168
 QY 165 FGGLTGAIAELIP-EIKGLEFALCSLFTLLDSCRTKKQIPSLLAGLSFTLIV 223
 DB 169 LGTVAGYTGALIPDKTFGLDAIFPALILITSLALKNK---VTRKAAFPAGSTLALITT 225
 QY 224 PGQALFAALLIFL-GLT 239
 DB 226 PELASGLPLISLFLGI 242

RESULT 11

US-09-134-000C-6789
 Sequence 6789, Application US/09134000C
 Patent No. 6617156
 GENERAL INFORMATION:
 APPLICANT: Lynn Doucette-Stamm et al
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
 FILE REFERENCE: 032796-032
 CURRENT APPLICATION NUMBER: US/09/134,000C
 CURRENT FILING DATE: 1998-08-13
 PRIOR APPLICATION NUMBER: US 60/055,778
 PRIOR FILING DATE: 1997-08-15
 NUMBER OF SEQ ID NOS: 6812
 SEQ ID NO 6789
 SOFTWARE: PatentIn version 3.1
 SEQ ID NO 6789
 LENGTH: 206
 TYPE: PRT
 ORGANISM: Enterococcus faecalis
 US-09-134-000C-6789

Query Match 10.8%; Score 137; DB 4; Length 206;
 Best Local Similarity 26.8%; Pred. No. 1e-07; Indels 24; Gaps 9;
 Matches 52; Conservative 42; Mismatches 76;

QY 71 SGLIFAGSTEMLVIALVGAAPLGAIALTTLLVNRHVFYAFSF-PLHVNKNPIARFYSV 129
 DB 17 SFPIFAGSAQFTVSMILGSGPSLISVLATFLVNAKMIIMGMTIAPYFRABSLGRMLWG 76
 QY 130 FALIDEAYAV-----TAAAP--GMSAMRLISMOIAFHSYVWPGGLTGAIAELIP- 179
 DB 77 TLLTDESFLGKMKLNHTGRUSFEWFNANLIS-----YAVWFTTIGAVLGRFIANP 131
 QY 180 EIKGLEFALCSLIV--TLLDSCRTKKQIPSLLAGLSF--TIALVPIGOAL-FAAL 232

Query Match	8.9%;	Score 113.5;	DB 4;	Length 484;
Best Local Similarity	25.1%;	Pred. No. 0.00018;		

Sequence 31521, Application US/09252991A

```
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 31521
; LENGTH: 555
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
; US-09-252-991A-31521

Query Match      8.2%; Score 104; DB 4; Length 555;
Best Local Similarity 22.1%; Pred. No. 0.0028;
Matches 75; Conservative 32; Mismatches 92; Indels 140; Gaps 16;

QY 32 IAQGLKTSIAAGLGMVPI---GIAFGLLVYQYGVEMWAAFLFSGL-----IFAGSTEML 82
Db 216 ICMGLVALLLPWIGRWLAERHGIYCGLLAIAAC-IMRELDGGLALIASALAGSGVAI 274
QY 83 VIALVGA-----APGAIATLT 100
Db 275 IQALVPEGVVGRVPRVPAAMGLYSASIMAGCGTAATLSPRIAEHPSNMQAGANAVPA 334
QY 101 LVNFRHVFYAFSFPPLHV---KNPIARFYS-----VPAIDEAIVATA----- 142
Db 335 LLA---LLWMFARPREVLPSAGEPVRIHFFGNRGWLLAVYFGLINGYTSWVAMLPHY 391
QY 143 -RPAGWSA-----WRLISWQIAFHSYWFVGG 168
Db 392 YRQLGWSRQDSGGVIGMTITFQVLAALSVPLLIRRLDRRPMILALVQ-----LGGF 445
QY 169 TGVAIABEIPPEIKGLEFAL-----C-SLEFVTLTDSCKTKQIPSL--LLAGLSFTI 218
Db 446 CGLL---LMPQHAALWVALIGYGCFAFSLTTLTDLHLEPRPAGSLAAFGIGIFII 502
QY 219 ALVY-----IPGQ-----ALFALLIFGLITRY 243
Db 503 TGIVPYLTGMRLDVTGSGFQASWMLLAASVAMLVTLRF 541
```

Search completed: April 16, 2004, 06:56:34
Job time : 24.7716 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 16, 2004, 06:53:52 ; Search time 17.4791 Seconds
(without alignments)
386.360 Million cell updates/sec

Title: US-10-608-504-3
Perfect score: 1269
Sequence: 1 MKNQEHSHSLVSPSKAAL.....LIFLGLTLTYFLGKAK 251

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 193692 seqs, 26905285 residues

Total number of hits satisfying chosen parameters: 193692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_AA_New.*
1: /cgn2_6/prodata/1/paa/PCT_NEW_COMB.pep.*
2: /cgn2_6/prodata/1/paa/US06_NEW_COMB.pep.*
3: /cgn2_6/prodata/1/paa/US07_NEW_COMB.pep.*
4: /cgn2_6/prodata/1/paa/US08_NEW_COMB.pep.*
5: /cgn2_6/prodata/1/paa/US09_NEW_COMB.pep.*
6: /cgn2_6/prodata/1/paa/US10_NEW_COMB.pep.*
7: /cgn2_6/prodata/1/paa/US60_NEW_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	185.5	14.6	255	US-10-417-884A-4920	Sequence 4920, Ap
2	161.5	12.7	242	US-10-724-972A-6721	Sequence 6721, Ap
3	99	7.8	455	US-60-556-841-2480	Sequence 2480, Ap
4	98	7.7	477	US-60-556-841-7285	Sequence 7285, Ap
5	98	7.7	760	PCT-US04-03417-32	Sequence 32, Appl
6	98	7.7	760	US-10-772-636-32	Sequence 32, Appl
7	98	7.7	858	US-10-725-081-14	Sequence 14, Appl
8	98	7.7	858	US-60-554-751-14	Sequence 14, Appl
9	96.5	7.6	417	US-09-857-097C-148	Sequence 148, Appl
10	96.5	7.6	417	US-09-857-097C-149	Sequence 149, Appl
11	96.5	7.6	417	US-09-857-097C-150	Sequence 150, Appl
12	96.5	7.6	417	US-09-857-097C-152	Sequence 152, Appl
13	92.5	7.3	366	US-10-417-884A-3665	Sequence 3665, Ap
14	91.5	7.2	389	US-60-556-841-2463	Sequence 2463, Ap
15	90	7.1	417	US-09-857-097C-151	Sequence 151, Appl
16	90	7.1	560	US-60-556-841-2332	Sequence 2332, Ap
17	89.5	7.1	552	US-10-781-014-654	Sequence 654, Appl
18	88.5	7.0	144	US-10-767-701-61776	Sequence 61776, A
19	88	6.9	471	US-60-556-841-49557	Sequence 4957, Ap
20	88	6.9	476	US-60-556-841-12104	Sequence 12104, A
21	87.5	6.9	492	US-10-724-972A-6373	Sequence 6373, Ap
22	87	6.9	497	US-60-556-841-10192	Sequence 10192, A
23	87	6.9	508	US-60-556-841-6098	Sequence 6098, Ap
24	87	6.9	555	US-10-100-683-7076	Sequence 7076, Ap
25	86.5	6.8	306	US-10-275-027A-322	Sequence 322, Appl
26	86	6.8	850	US-10-725-081-4	Sequence 4, Appl

27	86	6.8	852	US-60-554-751-12	Sequence 12, Appl
28	85.5	6.7	235	US-10-417-884A-3775	Sequence 3775, Ap
29	85.5	6.7	476	US-60-556-841-8474	Sequence 8474, Ap
30	85	6.7	392	US-60-556-841-4013	Sequence 4013, Ap
31	84.5	6.7	388	PCT-US04-07467-564	Sequence 564, Appl
32	84.5	6.7	426	US-60-556-841-6894	Sequence 6894, Ap
33	84	6.6	471	US-60-556-841-3188	Sequence 3188, Ap
34	84	6.6	858	US-60-554-751-13	Sequence 13, Appl
35	83.5	6.6	382	US-60-556-841-403	Sequence 403, Appl
36	83.5	6.6	453	US-60-556-841-6300	Sequence 6300, Ap
37	83	6.5	465	US-60-556-841-8829	Sequence 8829, Ap
38	82.5	6.5	305	US-10-417-884A-4215	Sequence 4215, Ap
39	82.5	6.5	375	US-10-767-701-43678	Sequence 43678, A
40	82	6.5	304	US-60-556-841-1622	Sequence 1622, Ap
41	82	6.5	454	US-60-556-841-2905	Sequence 2905, Ap
42	82	6.5	464	US-60-556-841-8095	Sequence 8095, Ap
43	82	6.5	469	US-60-556-841-11213	Sequence 11213, A
44	82	6.5	634	US-09-830-915D-37	Sequence 37, Appl
45	82	6.5	675	US-60-551-121-14	Sequence 14, Appl

ALIGNMENTS

RESULT 1
US-10-417-884A-4920
; Sequence 4920, Application US/1041784A
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/417, 884A
; FILING DATE: 17-Apr-2003
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/107, 532
; FILING DATE: 30-Jun-1998
; APPLICATION NUMBER: 60/085, 598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Arinello, Pamela Deneke
; REGISTRATION NUMBER: 40, 489
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 4920:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 255 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...255

SEQUENCE DESCRIPTION: SEQ ID NO: 4920:
US-10-417-884A-4920

Query Match 14.6%; Score 185.5; DB 6; Length 255;
Best Local Similarity 25.0%; Pred. No. 8.1e-12;
Matches 63; Conservative 51; Mismatches 105; Indels 33; Gaps 10;

16 SKALBDDKGYRREYIAQGLKTSIAAGCMYPIGIAFGILVYQYEWMAAPLPSGLIF 75
15 AKGVLLIDK---LDKTKAIKOTLPFTVFGYIGLGFVGAAGHPHLVTLSSLVY 70
76 AGSTEMLVIALVYGAAPLGAIALTLTLVNERHVFYAFSPFLHVKNPPIARFYSVPALI-D 134
71 AGSAQFTVWMLASHPSPLSLVSTPLVNSRMILMSWTIAPYFKKRLQLNLLIGTLTD 130
135 EAYAVTPARPAAGSAMRLISMOIAFH-----SY-WYFGSLTGAIAELIPEIK-G 183
131 EEPAL-----GNNKLVNTGQKLNFRWMTANNISYLVWSSLVGALLGNFTDPRKEG 184
184 LEFALCSLFTL-----TLDSCKRTKQIPSLLAGLFT--IALVYIPOGALFALLIF- 235
185 LDRIVAMFTGLLYQVYISDRNRSKR-QLILIGTLILVYGLIIFPSNLVYVVVTLIG 243
236 --LGLLTIRYF 245
244 CGLGVWIKHAF 255

RESULT 2

US-10-724-972A-6721
Sequence 6721, Application US/10724972A
GENERAL INFORMATION:
APPLICANT: Doucette-Stamm, Lynn
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
FILE REFERENCE: PATHO3-16
CURRENT APPLICATION NUMBER: US/10/724, 972A
CURRENT FILING DATE: 2003-12-01
PRIOR APPLICATION NUMBER: 09/450, 969
PRIOR FILING DATE: 1999-11-29
PRIOR APPLICATION NUMBER: 09/134, 001
PRIOR FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: 60/064, 964
PRIOR FILING DATE: 1997-11-08
PRIOR APPLICATION NUMBER: 60/055, 779
PRIOR FILING DATE: 1997-08-14
NUMBER OF SEQ ID NOS: 7544
SEQ ID NO 6721
LENGTH: 242
TYPE: PRT
ORGANISM: S. epidermidis
US-10-724-972A-6721

Query Match 12.7%; Score 161.5; DB 6; Length 242;
Best Local Similarity 24.8%; Pred. No. 2.4e-09;
Matches 57; Conservative 40; Mismatches 90; Indels 43; Gaps 8;

34 QGLKTSIAAGCMYPIGIAFGILVYQYEWMAAPLPSGLIFAGSTEMLVIALVYGAAPL 93
20 QGVCECPITLIGVAGVLSFGIYASQNSFVLEITLCLITVYGAQFICTIVAGTPI 79
94 GATLTLTLVNERHVFYAFSPFLHVKNPPIARFYSVP-----ALDEAVATAARP-- 144
80 SAIVLTLLVNSR-----MFLSLMTLAPYKQGFENRVLGTLTLDERFGV-ALTTPYV 132
145 -----AGNSAMRLISMOIAFHSHYVWFGSLTGAIAELIPEI-EIKGLEFALCSLFTVLD 197
133 KGEKINDR--LHGINTAVYLFWTVSCVIGALFGYISINPALGIDFAITGMFTFLCI 188
198 DSCRTKQ-----IPSLLAGLFTIALVYIPOGALFALL 233
189 SQFGIKSKRLRIYIVLVICVIMLLSLPEYVAIIL--AIVTALL 236

RESULT 3

US-60-556-841-2490
Sequence 2490, Application US/60556841
GENERAL INFORMATION:
APPLICANT: Abad, Mark S.
TITLE OF INVENTION: Genes and Uses for Plant Improvement
FILE REFERENCE: 38-21(53450)
CURRENT APPLICATION NUMBER: US/60/556, 841
CURRENT FILING DATE: 2004-03-25
NUMBER OF SEQ ID NOS: 12463
SEQ ID NO 2490
LENGTH: 455
TYPE: PRT
ORGANISM: Bacillus halodurans
US-60-556-841-2490

Query Match 7.8%; Score 99; DB 7; Length 455;
Best Local Similarity 27.5%; Pred. No. 0.017;
Matches 56; Conservative 26; Mismatches 68; Indels 54; Gaps 11;

45 GMYPIGIAFGILVYQYEWMAAPLPSGLIFAGSTEMLVIALVY---GAAPLGAIALTT 100
191 GLFPEGF-LGL-----WPAIFGFYAFGIEIGLILANRLEEGDAKSGRAML-- 238
101 LVNERHVFYAFSPFLHVKNPPIARFYSVPALIDEAVATTAAPAGSAMRLISMOIAFH 160
239 LLTFR---LVVVSGLAILVPLNRFT-----ETSPFIATKRYG-----LPIVSH 282
161 SYWVFGSLTGAIAELIPEIKGLEFALCSLFTVTLDSCKRTKQIPSLLAGLS 215
283 --LFGNGLIILAGSSTWASL-----FAVTSILVTLAAD-----KQAPNLFARKVK 325
216 FTIALVYIPOGALFALLIFLGL 239
326 NKIALPALGLNIGMALSLVSL 349

RESULT 4

US-60-556-841-7285
Sequence 7285, Application US/60556841
GENERAL INFORMATION:
APPLICANT: Abad, Mark S.
TITLE OF INVENTION: Genes and Uses for Plant Improvement
FILE REFERENCE: 38-21(53450)
CURRENT APPLICATION NUMBER: US/60/556, 841
CURRENT FILING DATE: 2004-03-25
NUMBER OF SEQ ID NOS: 12463
SEQ ID NO 7285
LENGTH: 477
TYPE: PRT
ORGANISM: Ferrioplasma acidarmanus
US-60-556-841-7285

Query Match 7.7%; Score 98; DB 7; Length 477;
Best Local Similarity 24.6%; Pred. No. 0.023;
Matches 55; Conservative 31; Mismatches 86; Indels 52; Gaps 10;

53 FGLVYQYEWMAAPLPSGLIFAGSTEMLVIALVYGA---APLGAIALTT 101
92 YGLII---GYMWSYFVITNALIFTSQVIGIPGITSYFLHFNGSYTWIPI---MTAF 144
102 LVNERHVFY-----AFSPFLHVKNPPIARFYSVPALI---DEAVATAAPAGSAM 150
145 IIVLVIALVYLGKPSLAFSTASVIBIALVITSVIIVALGSKNTVPTFAGADGPA 204
151 RLISMOIAFHSHYVWFGSLTGAIAELIPEIKGLEFALCSLFTV-- 194
205 VAVGVNLAFIS--MSGSAVVTLGEEAQPRKNIKALLSFIITGVFVLTSTALTVGW 262
195 -TLDSCKRTKQIPSLLAGLFTIALVYIPOGALFALL 233

Tue Apr 20 06:47:18 2004

us-10-608-504-3.rapn

Page 3

Db 263 GVSXMSFTSTSPGLIVSDKYLGPFTIVLFVPIINSIFAGSL 306

RESULT 5
PCT-US04-03417-32
Sequence 32, Application PC/TUS0403417
GENERAL INFORMATION:
APPLICANT: Kelly, Louise M.
APPLICANT: Carroll, Joseph M.
APPLICANT: Farlow, Deborah
APPLICANT: Healy, Aileen
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING
TITLE OF INVENTION: HEMATOLOGICAL DISORDERS USING 9118, 990, 17662, 81982, 630,
TITLE OF INVENTION: 21472, 17692, 19290, 21620, 21689, 28899, 53659, 64549,
TITLE OF INVENTION: 9465, 23544, 7366, 27417, 57259, 21844, 943, 2061, 5891,
TITLE OF INVENTION: 9137, 13908, 14310, 17600, 25584, 27824, 28469, 38947,
TITLE OF INVENTION: 53003, 965, 56639, 9661, 16052, 1521, 6662, 13913, 12405 OR
FILE REFERENCE: MPI03-015
CURRENT APPLICATION NUMBER: PCT/US04/03417
CURRENT FILING DATE: 2004-02-05
PRIOR APPLICATION NUMBER: US 60/445,241
PRIOR FILING DATE: 2003-02-05
PRIOR APPLICATION NUMBER: US 60/448,389
PRIOR FILING DATE: 2003-02-18
PRIOR APPLICATION NUMBER: US 60/456,320
PRIOR FILING DATE: 2003-03-20
PRIOR APPLICATION NUMBER: US 60/460,279
PRIOR FILING DATE: 2003-04-03
PRIOR APPLICATION NUMBER: US 60/465,924
PRIOR FILING DATE: 2003-04-28
PRIOR APPLICATION NUMBER: US 60/470,052
PRIOR FILING DATE: 2003-05-13
PRIOR APPLICATION NUMBER: US 60/498,106
PRIOR FILING DATE: 2003-08-26
PRIOR APPLICATION NUMBER: US 60/500,179
PRIOR FILING DATE: 2003-09-04
PRIOR APPLICATION NUMBER: US 60/502,909
PRIOR FILING DATE: 2003-09-15
PRIOR APPLICATION NUMBER: US 60/510,351
PRIOR FILING DATE: 2003-10-10
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 80
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 32
LENGTH: 760
TYPE: PRT
ORGANISM: Homo sapiens
PCT-US04-03417-32

Query Match 7.7%; Score 98; DB 1; Length 760;
Best Local Similarity 25.6%; Pred. No. 0.042;
Matches 40; Conservative 27; Mismatches 55; Indels 34; Gaps 8;

Db 80 EMVLIVLVG-----AAPGAIATLLVNRHVFAFSPPLHVKNPIARFVSVALID 134
254 EVLSAAAAGVSAFGAPIGV-----LFSLEVSX--YFPLKTL-----WRSFFALV 300

Qy 135 EAYAVTARPAAGSAMRLISMOIAFHSYVWFGJTGVAIAELIPEIKGLEFALC-SLFV 193
Db 301 AAFTRLSINPFGNS-RLVLFYVEYHTPMY-----MAELPFIILGVFGJLWGLTFLI 350

Qy 194 TLTLDSCTKX-----QIPSLLAGLSFTTALVVP 224
Db 351 RCNIAMCRKRTKTRLGKXPVLEVIIVTATITAYP 386

RESULT 6
US-10-772-636-32
Sequence 32, Application US/10772636
GENERAL INFORMATION:
APPLICANT: Kelly, Louise M.
APPLICANT: Carroll, Joseph M.

APPLICANT: Farlow, Deborah
APPLICANT: Healy, Aileen
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING
TITLE OF INVENTION: HEMATOLOGICAL DISORDERS USING 9118, 990, 17662, 81982, 630,
TITLE OF INVENTION: 21472, 17692, 19290, 21620, 21689, 28899, 53659, 64549,
TITLE OF INVENTION: 9465, 23544, 7366, 27417, 57259, 21844, 943, 2061, 5891,
TITLE OF INVENTION: 9137, 13908, 14310, 17600, 25584, 27824, 28469, 38947,
TITLE OF INVENTION: 53003, 965, 56639, 9661, 16052, 1521, 6662, 13913, 12405 OR
FILE REFERENCE: MPI03-015P1NOMNIM
CURRENT APPLICATION NUMBER: US/10/772,636
CURRENT FILING DATE: 2004-02-05
PRIOR APPLICATION NUMBER: US 60/445,241
PRIOR FILING DATE: 2003-02-05
PRIOR APPLICATION NUMBER: US 60/448,389
PRIOR FILING DATE: 2003-02-18
PRIOR APPLICATION NUMBER: US 60/456,320
PRIOR FILING DATE: 2003-03-20
PRIOR APPLICATION NUMBER: US 60/460,279
PRIOR FILING DATE: 2003-04-03
PRIOR APPLICATION NUMBER: US 60/465,924
PRIOR FILING DATE: 2003-04-28
PRIOR APPLICATION NUMBER: US 60/470,052
PRIOR FILING DATE: 2003-05-13
PRIOR APPLICATION NUMBER: US 60/498,106
PRIOR FILING DATE: 2003-08-26
PRIOR APPLICATION NUMBER: US 60/500,179
PRIOR FILING DATE: 2003-09-04
PRIOR APPLICATION NUMBER: US 60/502,909
PRIOR FILING DATE: 2003-09-15
PRIOR APPLICATION NUMBER: US 60/510,351
PRIOR FILING DATE: 2003-10-10
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 80
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 32
LENGTH: 760
TYPE: PRT
ORGANISM: Homo sapiens
US-10-772-636-32

Query Match 7.7%; Score 98; DB 6; Length 760;
Best Local Similarity 25.6%; Pred. No. 0.042;
Matches 40; Conservative 27; Mismatches 55; Indels 34; Gaps 8;

Qy 80 EMVLIVLVG-----AAPGAIATLLVNRHVFAFSPPLHVKNPIARFVSVALID 134
Db 254 EVLSAAAAGVSAFGAPIGV-----LFSLEVSX--YFPLKTL-----WRSFFALV 300

Qy 135 EAYAVTARPAAGSAMRLISMOIAFHSYVWFGJTGVAIAELIPEIKGLEFALC-SLFV 193
Db 301 AAFTRLSINPFGNS-RLVLFYVEYHTPMY-----MAELPFIILGVFGJLWGLTFLI 350

Qy 194 TLTLDSCTKX-----QIPSLLAGLSFTTALVVP 224
Db 351 RCNIAMCRKRTKTRLGKXPVLEVIIVTATITAYP 386

RESULT 7
US-10-725-081-14
Sequence 14, Application US/10725081
GENERAL INFORMATION:
APPLICANT: ADLER, JON ELLIOT
APPLICANT: ZOULIYA, SERGEY
APPLICANT: LI, XINDONG SHAN
APPLICANT: O'CONNELL, SHAN
APPLICANT: STASZEWSKI, LENA
TITLE OF INVENTION: TIR TASTE RECEPTORS AND GENES ENCODING SAME
FILE REFERENCE: 078003/0277870/RXT
CURRENT APPLICATION NUMBER: US/10/725,081
CURRENT FILING DATE: 2003-12-02
PRIOR APPLICATION NUMBER: 60/187,546
PRIOR FILING DATE: 2000-03-07

PRIOR APPLICATION NUMBER: 60/195,536
 PRIOR FILING DATE: 2000-04-07
 PRIOR APPLICATION NUMBER: 60/209,840
 PRIOR FILING DATE: 2000-06-06
 PRIOR APPLICATION NUMBER: 60/214,213
 PRIOR FILING DATE: 2000-06-23
 PRIOR APPLICATION NUMBER: 30/226,448
 PRIOR FILING DATE: 2000-08-17
 PRIOR APPLICATION NUMBER: 60/259,227
 PRIOR FILING DATE: 2001-01-03
 NUMBER OF SEQ ID NOS: 20
 SOFTWARE: Patent In Ver. 2.1
 SEQ ID NO 14
 LENGTH: 858
 TYPE: PRT
 ORGANISM: Rattus sp.
 US-10-725-081-14

Query Match 7.7%; Score 98; DB 6; Length 858;
 Best Local Similarity 22.3%; Pred. No. 0.048;
 Matches 59; Conservative 35; Mismatches 92; Indels 78; Gaps 12;

52 AAGLLVQYGYEMWAPLPS-----GLIPAGSTEMLVY-----ALVGAAPLG 94
 591 ALGLFV-----HYWDSPLVQASGSLFCFGLICGLFCLSVLLPGRPRSASCLAQQPM 645
 95 AIALT-----TLVNFRRHVFYAFSPFL-----HVKNPDIAPRYSVFALIDE---AVAV 139
 646 HLPITGLSTLFLQAAEIFESELPLSMANWLCISYRGWAMLVLLATLVEALCAWYL 705
 140 TAARPAGSAMRLISMQIAFH---SYWVEGLTGV---ATAELIPF-----E 180
 706 MAPPEVYTMQVLPTEVELEHCHMRSMWSLGLVHINNAVLAFICFGLTFLVQSGPRNKR 765
 181 IKGLEPACSLFVTLTLDSCRTKQIPLSLAGLSFTIALVPIGQALPAL----- 232
 766 ARGITFAMLAYFLTW-----SFVP---LLANVQVAPVQVQMGAILFCALGIATFHL 816
 233 -----LIFGLTLTRFPLGKAK 251
 817 PKCYLMLPELNTQEFFLGSRPK 840

RESULT 8
 US-60-554-751-14
 Sequence 14, Application US/60554751
 GENERAL INFORMATION:
 APPLICANT: L1, Xia
 APPLICANT: L1, Weihua
 APPLICANT: Reed, Danielle R.
 APPLICANT: Bachmanov, Alexander A.
 APPLICANT: Brand, Joseph G.
 TITLE OF INVENTION: TASTE RECEPTOR OF THE T1R FAMILY FROM DOMESTIC CAT
 FILE REFERENCE: NON-0298
 CURRENT APPLICATION NUMBER: US/60/554,751
 CURRENT FILING DATE: 2004-03-19
 NUMBER OF SEQ ID NOS: 98
 SOFTWARE: Patent In version 3.2
 SEQ ID NO 14
 LENGTH: 858
 TYPE: PRT
 ORGANISM: Rattus rattus
 US-60-554-751-14

Query Match 7.7%; Score 98; DB 7; Length 858;
 Best Local Similarity 22.3%; Pred. No. 0.048;
 Matches 59; Conservative 35; Mismatches 92; Indels 78; Gaps 12;

52 AAGLLVQYGYEMWAPLPS-----GLIPAGSTEMLVY-----ALVGAAPLG 94
 591 ALGLFV-----HYWDSPLVQASGSLFCFGLICGLFCLSVLLPGRPRSASCLAQQPM 645
 95 AIALT-----TLVNFRRHVFYAFSPFL-----HVKNPDIAPRYSVFALIDE---AVAV 139

646 HLPITGLSTLFLQAAEIFESELPLSMANWLCISYRGWAMLVLLATLVEALCAWYL 705
 140 TAARPAGSAMRLISMQIAFH---SYWVEGLTGV---ATAELIPF-----E 180
 706 MAPPEVYTMQVLPTEVELEHCHMRSMWSLGLVHINNAVLAFICFGLTFLVQSGPRNKR 765
 181 IKGLEPACSLFVTLTLDSCRTKQIPLSLAGLSFTIALVPIGQALPAL----- 232
 766 ARGITFAMLAYFLTW-----SFVP---LLANVQVAPVQVQMGAILFCALGIATFHL 816
 233 -----LIFGLTLTRFPLGKAK 251
 817 PKCYLMLPELNTQEFFLGSRPK 840

RESULT 9
 US-09-857-097C-148
 Sequence 148, Application US/09857097C
 GENERAL INFORMATION:
 APPLICANT: Urbaniak, Stanislaw J.
 APPLICANT: Barker, Robert, N.
 TITLE OF INVENTION: ALLO- AND AUTO-REACTIVE T-CELL EPITOPES
 FILE REFERENCE: P097
 CURRENT APPLICATION NUMBER: US/09/857,097C
 CURRENT FILING DATE: 2001-05-31
 PRIOR APPLICATION NUMBER: 9826378.3
 PRIOR FILING DATE: 1998-12-01
 NUMBER OF SEQ ID NOS: 152
 SOFTWARE: Patent In Ver. 2.1
 SEQ ID NO 148
 LENGTH: 417
 TYPE: PRT
 ORGANISM: Homo sapiens
 FEATURE:
 OTHER INFORMATION: Rhce Residues 111-125
 US-09-857-097C-148

Query Match 7.6%; Score 96.5; DB 5; Length 417;
 Best Local Similarity 21.7%; Pred. No. 0.028;
 Matches 65; Conservative 41; Mismatches 90; Indels 103; Gaps 16;

18 AALEPDDKG-YRREYIAQGLKTSIAGLCMT-----IGIAFGLLVQYGYEMWAP 68
 36 ASLE-DQKGLVASYQVGDLTVMAGIGLFTSSFRHSWSVAENLMTALGVQW--AI 92
 69 LPSC-----LIPAGST-----EMLVIALVWGAAPLGA 95
 93 LLDGLSPGSKVITLFSRLATMSAMSVLISAGAVLGKVNLAQVWVLTALGT 152
 96 IAL-----TLVNFRRHVFYAFS--FPLHV---VKNPI-----ARFYSVALI 133
 133 LRMVSNIFNDYHNNLRH--FYVFAAYFGLVAMCDEKPEPKGTEDNDORATTPSLAML 211
 134 DEAYAVTARAPAGSAMRLISMQI-----AFSYWVFGGLTGVATAELIPFELK 183
 212 GALEL-----WMPFVSNSPLRSPQKMAENITY-----ALAVSVYVTLISG 255
 184 LEFALCSLFTVTLTLDSCRTKQIPLSLA--GLSFTIALVPIGQALPALIIFGLIT 241
 256 SELAPQRKISMT-----YHSAVLAVGAVAVGTSCHLIPSLAMVLGLVAGLISI 306

RESULT 10
 US-09-857-097C-149
 Sequence 149, Application US/09857097C
 GENERAL INFORMATION:
 APPLICANT: Urbaniak, Stanislaw J.
 APPLICANT: Barker, Robert, N.
 TITLE OF INVENTION: ALLO- AND AUTO-REACTIVE T-CELL EPITOPES
 FILE REFERENCE: P097
 CURRENT APPLICATION NUMBER: US/09/857,097C
 CURRENT FILING DATE: 2001-05-31

PRIOR APPLICATION NUMBER: 9826378.3
PRIOR FILING DATE: 1998-12-01
NUMBER OF SEQ ID NOS: 152
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 149
LENGTH: 417
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: Rhce Residues 121-135
US-09-857-097C-149

Query Match
Best Local Similarity 21.7%; Pred. No. 0.028; Length 417;
Matches 65; Conservative 41; Mismatches 90; Indels 103; Gaps 16;

QY 18 ALEPDDKG-YRREIAQGLKTSIAAGLGNTP-----IGIAFGLVIOGYEWMAAP 68
DB 36 ASLE-DQKGLVASVQVQDGLTWMAALGLGLTSSFRHSWSSVAFVFLFMLALGVQW--AI 92
QY 69 LFSG-----LIFAGST-----EMLVIALVVGAAPIGA 95
DB 93 LLDGFLSQPSPGKVITLFSIRLATWSAMSVLISAGAVLGKVNLAQLVWVWLVETALGT 152
QY 96 IAL-----TTLVNFRRHVFYAFS--FPLHV--VKNPI-----ARFVSVPALI 133
DB 153 LRMVINSINFTDYHNNLRH-FYVFAAVFGLTVAMCLPKPLPKGTEDNDQRAIIPSLAML 211
QY 134 DEAVAVTAARPAGMSAMRLISMQI-----AFHSYVWFGGLTGVAIAELIPEIKG 183
DB 212 GALFL-----WMFWSVNSALRLSPQKRNAMFNTY-----ALAVSVTAISG 255
QY 184 LEFALCSLFTVLTLDSCRTKKQIPSLIA-GLSFTIALVVIPOALFALLIFGLLTI 241
DB 256 SSLAHPOKISMT-----YHSAVLAGVAVGTSCHLIPSPMLAMVLGVAGLISI 306

RESULT 11

US-09-857-097C-150
Sequence 150, Application US/09857097C
GENERAL INFORMATION:
APPLICANT: Urbanek, Stanislaw J.
FILE REFERENCE: 0097
CURRENT APPLICATION NUMBER: US/09/857,097C
PRIOR FILING DATE: 2001-05-31
PRIOR APPLICATION NUMBER: 9826378.3
NUMBER OF SEQ ID NOS: 152
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 150
LENGTH: 417
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: Rhce Residues 131-145
US-09-857-097C-150

Query Match
Best Local Similarity 7.6%; Score 96.5; DB 5; Length 417;
Matches 65; Conservative 41; Mismatches 90; Indels 103; Gaps 16;

QY 18 ALEPDDKG-YRREIAQGLKTSIAAGLGNTP-----IGIAFGLVIOGYEWMAAP 68
DB 36 ASLE-DQKGLVASVQVQDGLTWMAALGLGLTSSFRHSWSSVAFVFLFMLALGVQW--AI 92
QY 69 LFSG-----LIFAGST-----EMLVIALVVGAAPIGA 95
DB 93 LLDGFLSQPSPGKVITLFSIRLATWSAMSVLISAGAVLGKVNLAQLVWVWLVETALGT 152
QY 96 IAL-----TTLVNFRRHVFYAFS--FPLHV--VKNPI-----ARFVSVPALI 133

DB 153 LRMVINSINFTDYHNNLRH-FYVFAAVFGLTVAMCLPKPLPKGTEDNDQRAIIPSLAML 211
QY 134 DEAVAVTAARPAGMSAMRLISMQI-----AFHSYVWFGGLTGVAIAELIPEIKG 183
DB 212 GALFL-----WMFWSVNSALRLSPQKRNAMFNTY-----ALAVSVTAISG 255
QY 184 LEFALCSLFTVLTLDSCRTKKQIPSLIA-GLSFTIALVVIPOALFALLIFGLLTI 241
DB 256 SSLAHPOKISMT-----YHSAVLAGVAVGTSCHLIPSPMLAMVLGVAGLISI 306

RESULT 12

US-09-857-097C-152
Sequence 152, Application US/09857097C
GENERAL INFORMATION:
APPLICANT: Urbanek, Stanislaw J.
FILE REFERENCE: 0097
CURRENT APPLICATION NUMBER: US/09/857,097C
PRIOR FILING DATE: 2001-05-31
PRIOR APPLICATION NUMBER: 9826378.3
NUMBER OF SEQ ID NOS: 152
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 152
LENGTH: 417
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: Rhce Residues 151-165
US-09-857-097C-152

Query Match
Best Local Similarity 21.7%; Pred. No. 0.028; Length 417;
Matches 65; Conservative 41; Mismatches 90; Indels 103; Gaps 16;

QY 18 ALEPDDKG-YRREIAQGLKTSIAAGLGNTP-----IGIAFGLVIOGYEWMAAP 68
DB 36 ASLE-DQKGLVASVQVQDGLTWMAALGLGLTSSFRHSWSSVAFVFLFMLALGVQW--AI 92
QY 69 LFSG-----LIFAGST-----EMLVIALVVGAAPIGA 95
DB 93 LLDGFLSQPSPGKVITLFSIRLATWSAMSVLISAGAVLGKVNLAQLVWVWLVETALGT 152
QY 96 IAL-----TTLVNFRRHVFYAFS--FPLHV--VKNPI-----ARFVSVPALI 133
DB 153 LRMVINSINFTDYHNNLRH-FYVFAAVFGLTVAMCLPKPLPKGTEDNDQRAIIPSLAML 211
QY 134 DEAVAVTAARPAGMSAMRLISMQI-----AFHSYVWFGGLTGVAIAELIPEIKG 183
DB 212 GALFL-----WMFWSVNSALRLSPQKRNAMFNTY-----ALAVSVTAISG 255
QY 184 LEFALCSLFTVLTLDSCRTKKQIPSLIA-GLSFTIALVVIPOALFALLIFGLLTI 241
DB 256 SSLAHPOKISMT-----YHSAVLAGVAVGTSCHLIPSPMLAMVLGVAGLISI 306

RESULT 13

US-10-417-884A-3665
Sequence 3665, Application US/10417884A
GENERAL INFORMATION:
APPLICANT: Lynn A Doucette-Stamm and David Bush
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
NUMBER OF SEQUENCES: 7310
CORRESPONDENCE ADDRESS:
ADDRESS: GENOME THERAPEUTICS CORPORATION
STREET: 100 Beaver Street
CITY: Waltham
STATE: Massachusetts
COUNTRY: USA
ZIP: 02354

COMPUTER READABLE FORM:

MEDIUM TYPE: CD/ROM ISO9660

COMPUTER: PC

OPERATING SYSTEM: <Unknown>

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/417,884A

FILING DATE: 17-Apr-2003

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/09/107,532

FILING DATE: 30-Jun-1998

APPLICATION NUMBER: 60/085,598

FILING DATE: 14 May 1998

APPLICATION NUMBER: 60/051571

FILING DATE: July 2, 1997

ATTORNEY/AGENT INFORMATION:

NAME: Ariniello, Pamela Deneke

REGISTRATION NUMBER: 40,489

REFERENCE/DOCKET NUMBER: GTC-012

TELECOMMUNICATION INFORMATION:

TELEPHONE: (781)893-5007

TELEFAX: (781)893-8277

INFORMATION FOR SEQ ID NO: 3665:

SEQUENCE CHARACTERISTICS:

LENGTH: 326 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

HYPOTHETICAL: YES

ORIGINAL SOURCE:

ORGANISM: Enterococcus faecium

FEATURE:

NAME/KEY: misc feature

LOCATION: (B) LOCATION 1...326

SEQUENCE DESCRIPTION: SEQ ID NO: 3665:

US-10-417-884A-3665

Query Match 7.3%; Score 92.5; DB 6; Length 326;

Best Local Similarity 24.6%; Pred. No. 0.055; Matches 34; Conservative 24; Mismatches 45; Indels 35; Gaps 6;

QY 143 RPAGWSAW-RLISMOIARHSYVFGLTGVAIAEIPFETKLEPAL-----CSLFV 193
 DB 6 RVDWWRGRVALVAILVLSISLFIQVSNVQLSEHLDQSQSLIILVTRLPRTICILV 65
 QY 134 TLTLDSC-----RRKKDIPSLLAG-----LSFTIALVVIPOAL-----FA 230
 DB 66 GATSSICGIMQHLTQNKFSPTTAGTMDSARLGIIVAMIFLPQASLIVRSLTAFCAFA 125
 QY 231 ALLIFGLITIRYFELGX 248
 DB 126 GTLLFSLTRL--FLQK 140

RESULT 14

US-60-556-841-2463

Sequence 2463, Application US/60556841

GENERAL INFORMATION:

APPLICANT: Abad, Mark S.

TITLE OF INVENTION: Genes and Uses for Plant Improvement

FILE REFERENCE: 38-21(53450)

CURRENT APPLICATION NUMBER: US/60/556,841

CURRENT FILING DATE: 2004-03-25

NUMBER OF SEQ ID NOS: 12463

SEQ ID NO 2463

LENGTH: 389

TYPE: PRT

ORGANISM: Brassica rapa

US-60-556-841-2463

Query Match 7.2%; Score 91.5; DB 7; Length 389;

Best Local Similarity 23.3%; Pred. No. 0.087; Matches 49; Conservative 28; Mismatches 90; Indels 43; Gaps 8;

QY 43 GLGWYPIGIAFGLIVIOGYEWMW-----APLESGLIIFAGS--TEMVLVIALVGAAPLG 94

DB 176 GLALIVYSHLFTALV---GAEWMAQOLGOSIFLFSWPFVVAIQTSPRAVLYMMIAFA--- 229

QY 95 AIALTLINFRHVFVAFFPLHVNKPIARFVSFPALIDEAVVTAARPAWGASRLIS 154

DB 220 --VIPTAVINVSNTKVVQSRKSGMLALANLFFVVL-----GGLVIMSLIS 276

QY 155 MOIARHSY-----NVFGSLTGVAIAELIPEIKLEFALCS--LVTTLTDSCTK 203

DB 277 PINIETVPHLVYVLTGIAFGFLVGRMILALCEBPGLKTMCMSSLYLFFALANALTA 336

QY 204 KOIPSLLAGISFTALVVIPOALFPALL 233

DB 337 R-----LNAGVPLVDLWVLGVCIFVSL 361

RESULT 15

US-09-857-097C-151

Sequence 151, Application US/09857097C

GENERAL INFORMATION:

APPLICANT: Urbanjak, Stanislaw J.

TITLE OF INVENTION: ALLO- AND AUTO-REACTIVE T-CELL EPITOPES

FILE REFERENCE: P097

CURRENT APPLICATION NUMBER: US/09/857,097C

CURRENT FILING DATE: 2001-05-31

PRIOR APPLICATION NUMBER: 9826378.3

PRIOR FILING DATE: 1998-12-01

NUMBER OF SEQ ID NOS: 152

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 151

LENGTH: 417

TYPE: PRT

ORGANISM: Homo sapiens

FEATURE:

OTHER INFORMATION: Rhd Residues 141-155

US-09-857-097C-151

Query Match 7.1%; Score 90; DB 5; Length 417;

Best Local Similarity 21.5%; Pred. No. 0.14; Matches 69; Conservative 36; Mismatches 104; Indels 112; Gaps 13;

QY 18 AALEPDKG-YRRYEIAQGLKTSLAGLGYF-----ICIAFGLVIOGYEWMWAP 68

DB 36 ASLE-DQKGLVASVYQVQDQTLWMAAIGLGLTSSFRHSWSSVAFTLALGVQW--AI 92

QY 69 LFSGLIFAGSTEMLV-----ALVGAAPIGAIALTLVY-----N 104

DB 93 LLDGFLSQPSPGKVITLFSIRLATMSALSVLISDAVIGKNLAQLVVMVVEVTALCN 152

QY 105 FR-----HVFYAF----- 112

DB 153 LRNVINIFNTVDYHMMQHIVFAAYFGSVAMCLEPKPLPEGTEDNDQTATIPISAMLG 212

QY 113 -----SPLHVKNPRIARFVSFPALIDEAVVTAARPAWGASRLISNQ-----IA 158

DB 213 ALFLMTFWSFNSALRSPIERKNVFN--NTYAAVAVVTALISSSLAHQOKSKTY 269

QY 159 FHSYVWFGGLTGVAIAELIPFB-----IKGLEFALCSLFTLTDSCTK-KOIPSLLAG 213

DB 270 VHSANVLAGVAVGTSCHLIPSPWLAVLGLVAGLVGAKVLPGCCNVLGIPHSING 329

QY 214 LSFTIALVVIPOALFPALLI 234

DB 330 YNF--SLIGLIGELIYIVLV 348

Search completed: April 16, 2004, 07:03:18

Job time : 18.4791 secs

Db 555 GGATGATACCAAGAGTTTTCATTTGCTTGGAAATCCCAATTGAAGATTITTCAG 614
 QY 194 ATCTGAAGACTTAACTTTGGTCTCATTCGCTGGGATTAAGTGTGGCCATCTTC 253
 Db 615 CTCAGAGCAGATATATGTCATGCTGTATTAAGCCGGGTTCATTTTTCGACTT 674
 QY 254 TTGGCGGTGACGACCTTGTGAGCGCTGGCGCTGG 290
 Db 675 TTTTACTTGCTTTGATGCTGAAGCTAGCTGGGG 711

RESULT 15

CK295424

LOCUS

DEFINITION

CK295424 936 bp mRNA linear EST 15-DEC-2003
 EST758138 Nicotiana benthamiana mixed tissue cDNA library,
 normalized, full-length Nicotiana benthamiana cDNA clone NBMCU77 5',
 end, mRNA sequence.

ACCESSION CK295424 GI:39879801
 VERSION CK295424.1
 KEYWORDS EST

SOURCE Nicotiana benthamiana
 ORGANISM Nicotiana benthamiana

REFERENCE 1 (bases 1 to 936)
 Buell, C.R., Hart, A., Zisman, V., Karamycheva, S.A., Day, B.,
 Staskawicz, B., Jin, H. and Baker, B.

TITLE Generation of EST sequences from Nicotiana benthamiana
 JOURNAL Unpublished (2003)
 COMMENT Other ESTs: EST758139

Contact: Robin Buell
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850, USA
 Email: potato-array@tigr.org
 Clones can be requested from TIGR via potato@tigr.org
 Seq primer: ATT TAG GTG ACA CTA TAG.

FEATURES
 source location/Qualifiers

1. 936
 /organism="Nicotiana benthamiana"
 /mol_type="mRNA"
 /db_xref="taxon:4100"
 /clone="NBMCU77"
 /tissue_type="abiotic and biotic stress-treated leaves,
 callus tissue and root tissue"
 /lab_host="MDH10B-Tona"
 /clone_lib="Nicotiana benthamiana mixed tissue cDNA
 library, normalized, full-length"
 /note="Vector: pCMVSPORT6.1; Site_1: EcoRI; Site_2: NotI;
 supplier: RNA was isolated from Nicotiana benthamiana
 tissues that include callus, roots from liquid culture
 grown plants, heat-stressed leaves (38 C, 3 hr and 6 hr),
 cold-stressed leaves (5 C 3 hr, 6hr), and pathogen
 challenged leaves (Pseudomonas syringae pv tomato 12 hr;
 Xanthomonas campestris pv campestris 12 hr, 18hr,
 Pseudomonas syringae pv phaseolicola 18hr, and Xanthomonas
 campestris pv vesicatoria 18hr). RNA was isolated from
 these tissues and pooled in approximately equal molar
 amounts."

ORIGIN

Query Match 10.7%; Score 34.6; DB 14; Length 936;
 Best Local Similarity 47.5%; Fred. No. 26;
 Matches 103; Conservative 0; Mismatches 114; Indels 0; Gaps 0;
 QY 74 TTCCGTTCTTAATCCTTAAGCCCTACGTGAATCAATTTGTGGCAAAATGGCGATGT 133
 Db 419 TTCTAATGTATATGCTAGCCCGAACAATCTGAAGTGGCGGGAATTCGGCGTGT 478
 QY 134 GGATGCCAGCAGGAATCCTTGGCATTTTGACCGCATCAAGCTTTGGCAGCATGCGATAG 193
 Db 479 GGATGATACCAAGAGTTTTCATTTGCTTGGAAATCCCAATTGAAGATTITTCAG 538

QY 194 ATCTGAAGACTTAACTTTGGTCTCATTCGCTGGGATTAAGTGTGGCCATCTTC 253
 Db 539 CTCAGAGCAGATATATGTCATGCTGTATTAAGCCGGGTTCATTTTTCGACTT 598
 QY 254 TTGGCGGTGACGACCTTGTGAGCGCTGGCGCTGG 290
 Db 599 TTTTACTTGCTTTGATGCTGAAGCTAGCTGGGG 635

Search completed: April 17, 2004, 20:01:49
 Job time : 1050.18 secs

QY 74 TTCCGTTCTTAATCCTTAGACCCCTACGTAATCAATTTGTGGCAAAATGGGATGT 133
DB 404 TTCTAATGTTAATCGGTGACCCGAAACATCTCAAGTGCGCGGGAATTCGGCGTGT 463
QY 134 GGATGCGACAGAGAAATCCTTGCCATTTTGACCGCATCAAGTTTGCGAGCAATGGCATAG 193
DB 464 GGATGATACCAAGAGTTTTCATTTGCTTGGAATTCCTCAATTAATGAAGTTTTCAGAG 523
QY 194 ATCTGAAGACTCTAACCTTTGGTCTCATTTGCGCTTGCGATTACAGTGTGGCGCATCTTC 253
DB 524 CTCAGAGCAAGATTAATGTCATGCGCTGTATAGCCGCGGTTGCATGTTTTCACACTT 583
QY 254 TTGGCGGTGACGACGACCTTTGTGAGCGTTGGCGCTGG 290
DB 584 TTTTACTTGCGCTTTTGATGCTGAAGCTAGGCTGGGG 620

RESULT 13
LOCUS CK288648
DEFINITION 890 bp mRNA linear EST 15-DEC-2003
normalised, full-length Nicotiana benthamiana cDNA clone NEMB102 5' end, mRNA sequence.

ACCESSION CK288648
VERSION CK288648
KEYWORDS
SOURCE
ORGANISM Nicotiana benthamiana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; lamiales; Solanales; Solanaceae; Nicotiana.
1 (bases 1 to 890)
Buell, C.R., Hart, A., Zismann, V., Karanymcheva, S.A., Day, B., Szaskawicz, B., Jin, H. and Baker, B.
Generation of EST sequences from Nicotiana benthamiana
Unpublished (2003)
Contact: Robin Buell
The Institute for Genomic Research
9712 Medical Center Dr, Rockville, MD 20850, USA
Email: potato-array@tigr.org
Clones can be requested from TIGR via potato@tigr.org
Seq primer: ATT TAG GTG ACA CTA TAG.
Location/Qualifiers

FEATURES
source
1..890
/organism="Nicotiana benthamiana"
/mol_type="mRNA"
/db_xref="taxon:4100"
/clone="NEMB102"
/tissue_type="abiotic and biotic stress-treated leaves, callus tissue and root tissue"
/lab_host="DH10B-TonA"
/clone_lib="Nicotiana benthamiana mixed tissue cDNA library, normalized, full-length"
/note="Vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI; supplier: RNA was isolated from Nicotiana benthamiana tissues that include callus, roots from liquid culture grown plants, heat-stressed leaves (38 C, 3 hr and 6 hr), cold-stressed leaves (5 C 3 hr, 6hr), and pathogen challenged leaves (Pseudomonas syringae pv tomato 12 hr; Xanthomonas campestris pv campestris 12 hr, 18hr; Pseudomonas syringae pv phaseolicola 18hr, and Xanthomonas campestris pv vesicatoria 18hr). RNA was isolated from these tissues and pooled in approximately equal molar amounts."

ORIGIN
Query Match 10.7%; Score 34.6; DB 14; Length 890;
Best Local Similarity 47.5%; Pred. No. 25;
Matches 103; Conservative 0; Mismatches 114; Indels 0; Gaps 0;

QY 74 TTCCGTTCTTAATCCTTAGACCCCTACGTAATCAATTTGTGGCAAAATGGGATGT 133
DB 543 TTCTAATGTTAATCGGTGACCCGAAACATCTCAAGTGCGCGGGAATTCGGCGTGT 602

QY 134 GGATGCGACAGAGAAATCCTTGCCATTTTGACCGCATCAAGCTTTCGAGCAATGGCATAG 193
DB 603 GGATGATACCAAGAGTTTTCATTTGCTTGGAATTCCTCAATTAATGAAGTTTTCAGAG 662
QY 194 ATCTGAAGACTCTAACCTTTGGTCTCATTTGCGCTTGCGATTACAGTGTGGCGCATCTTC 253
DB 663 CTCAGAGCAAGATTAATGTCATGCGCTGTATAGCCGCGGTTGCATGTTTTCACACTT 722
QY 254 TTGGCGGTGACGACGACCTTTGTGAGCGTTGGCGCTGG 290
DB 723 TTTTACTTGCGCTTTTGATGCTGAAGCTAGGCTGGGG 759

RESULT 14
LOCUS CK295854
DEFINITION 922 bp mRNA linear EST 15-DEC-2003
normalised, full-length Nicotiana benthamiana cDNA clone NEMC72 5' end, mRNA sequence.

ACCESSION CK295854
VERSION CK295854
KEYWORDS
SOURCE
ORGANISM Nicotiana benthamiana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; lamiales; Solanales; Solanaceae; Nicotiana.
1 (bases 1 to 922)
Buell, C.R., Hart, A., Zismann, V., Karanymcheva, S.A., Day, B., Szaskawicz, B., Jin, H. and Baker, B.
Generation of EST sequences from Nicotiana benthamiana
Unpublished (2003)
Other ESTs: EST758569
Contact: Robin Buell
The Institute for Genomic Research
9712 Medical Center Dr, Rockville, MD 20850, USA
Email: potato-array@tigr.org
Clones can be requested from TIGR via potato@tigr.org
Seq primer: ATT TAG GTG ACA CTA TAG.
Location/Qualifiers

FEATURES
source
1..922
/organism="Nicotiana benthamiana"
/mol_type="mRNA"
/db_xref="taxon:4100"
/clone="NEMC72"
/tissue_type="abiotic and biotic stress-treated leaves, callus tissue and root tissue"
/lab_host="DH10B-TonA"
/clone_lib="Nicotiana benthamiana mixed tissue cDNA library, normalized, full-length"
/note="Vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI; supplier: RNA was isolated from Nicotiana benthamiana tissues that include callus, roots from liquid culture grown plants, heat-stressed leaves (38 C, 3 hr and 6 hr), cold-stressed leaves (5 C 3 hr, 6hr), and pathogen challenged leaves (Pseudomonas syringae pv tomato 12 hr; Xanthomonas campestris pv campestris 12 hr, 18hr; Pseudomonas syringae pv phaseolicola 18hr, and Xanthomonas campestris pv vesicatoria 18hr). RNA was isolated from these tissues and pooled in approximately equal molar amounts."

ORIGIN
Query Match 10.7%; Score 34.6; DB 14; Length 922;
Best Local Similarity 47.5%; Pred. No. 26;
Matches 103; Conservative 0; Mismatches 114; Indels 0; Gaps 0;

QY 74 TTCCGTTCTTAATCCTTAGACCCCTACGTAATCAATTTGTGGCAAAATGGGATGT 133
DB 495 TTCTAATGTTAATCGGTGACCCGAAACATCTCAAGTGCGCGGGAATTCGGCGTGT 554
QY 134 GGATGCGACAGAGAAATCCTTGCCATTTTGACCGCATCAAGCTTTCGAGCAATGGCATAG 193

ORIGIN

/clone="CS0D1052YB18"
/issue="type="PLACENTA COT 25-NORMALIZED"
/clone.lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was placed with a NotI-oligo (dT)
primer. Five prime end enriched, double-strand cDNA was
digested with NotI and cloned into the NotI and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

Query Match 10.7%; Score 34.6; DB 13; Length 1201;
Best Local Similarity 39.9%; Pred. No. 25;
Matches 63; Conservative 25; Mismatches 70; Indels 0; Gaps 0;

QY 154 GCCATTTTGACCCGATCAACGTTTGGCAGCAATGATGATGAAAGCTTAACTTT 213
DB 943 GACTTCTTCMGCCCGCGCGGCGGCTGTGTGGGAGCGCCGCCCTTCGGAGT 1002
QY 214 GGTCTCATTCGCGTGGATTACAGTGTGGCGCATCTTCTTGGCGATGACGACCTTG 273
DB 1003 GSTKSGGTTSMGGCCMCTTKAMAGTCCGGVGHYTTTTTTTGTCTCCCMCMCCCTC 1062
QY 274 TTGAGCGTTGGCGCTGGACCATCGTTTGTGACT 311
DB 1063 CAGTATGATGAGGAGGAGTACATCTGCTGTATAGT 1100

RESULT 11

CE239541 413 bp mRNA linear EST 30-APR-2003
LOCUS RSH13B11 two-month-old roots from clone 'Beaupre' grown for 19 days

DEFINITION under restricted irrigation Populus balsamifera subsp. trichocarpa
x Populus deltoides cDNA 5', mRNA sequence.

ACCESSION CE239541 GI:30233030

VERSION EST.
KEYWORDS Populus balsamifera subsp. trichocarpa x Populus deltoides

SOURCE Populus balsamifera subsp. trichocarpa x Populus deltoides
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; euroids; Malpighiales; Salicaceae; Salicaceae; Populus.

REFERENCE 1 (bases 1 to 413)
AUTHORS Kohler, A., Delanelle, C., Martin, D. and Martin, F.
TITLE The poplar root transcriptome: analysis of 7000 expressed sequence
tags

JOURNAL Unpublished (2003)
COMMENT Contact: Martin FM
Equipe de Microbiologie Forestiere
Institut National de la Recherche Agronomique
Centre INRA de Nancy, 54280 Champenoux, France
Tel: +33 383 39 40 80
Fax: +33 383 39 40 69
Email: fmarlin@nancy.inra.fr
Insert Length: 413 Std Error: 0.00
Seq primer: Fornat 5' AAGCGCGCATGTGTGTATACC.

FEATURES

source
1. 413
/organism="Populus balsamifera subsp. trichocarpa x
Populus deltoides"
/mol_type="mRNA"
/cultivar="Beaupre/"
/db_xref="taxon:3695"
/db_xref="taxon:3695"
/clone.lib="two-month-old roots from clone 'Beaupre' grown
for 19 days under restricted irrigation"
/note="Organ: root; Vector: pTIP1EX2; cDNA library of
roots from two-month-old Populus trichocarpa Torr. & Gray
x deltoides Bartr. Ex Marshall (clone 'Beaupre') grown for
19 days under restricted irrigation to reach 50% of the
transpiration rate of fully watered plants. The cDNA
library was constructed from 1 ug of total RNA using the
SMART cDNA synthesis kit (Clontech, Palo Alto, CA, USA)
according to the manufacturer's instructions. The
resulting cDNA was packed into lambda phages using the

ORIGIN

Gigapack III Gold packaging kit (Stratagene, La Jolla,
CA). The pTIP1EX2 phagemid clones in Escherichia coli
were obtained by using the mass in vivo excision protocol
according to the manufacturer's instructions (Clontech)."

Query Match 10.7%; Score 34.6; DB 14; Length 413;
Best Local Similarity 60.4%; Pred. No. 19;
Matches 55; Conservative 0; Mismatches 36; Indels 0; Gaps 0;

QY 65 TCCGGCGGTTCCGTTCTTAATCCTTAAGCCCTTACCTGAATCAATTTGTGGCAAAA 124
DB 71 TCCAGAGGCTCCCTCTGAGACTATTNACCCCTCCAGGATTCACCTTCAGGTCAAAA 130
QY 125 TGGGATGTGATGCCAGCAAGATCTTCG 155
DB 131 TGGGATTCGGGCGCACAGAGATTCCTGCG 161

RESULT 12

CK297477 853 bp mRNA linear EST 15-DEC-2003
LOCUS EST760191 Nicotiana benthamiana mixed tissue cDNA library,
DEFINITION normalized, full-length Nicotiana benthamiana cDNA clone NEMD912 5'
end, mRNA sequence.

ACCESSION CK297477 GI:39883902
VERSION EST.
KEYWORDS Nicotiana benthamiana

SOURCE Nicotiana benthamiana
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanales; Solanales; Nicotiana.

REFERENCE 1 (bases 1 to 853)
AUTHORS Buell, C.R., Hart, A., Zismann, V., Karamycheva, S.A., Day, B.,
Stackawicz, B., Jin, H. and Baker, B.
TITLE Generation of EST sequences from Nicotiana benthamiana
JOURNAL Unpublished (2003)
COMMENT Other ESTs: EST760192
Contact: Robin Buell
The Institute for Genomic Research
9712 Medical Center Dr. Rockville, MD 20850, USA
Email: potato-array@tigr.org
Clones can be requested from TIGR via potato@tigr.org
Seq primer: ATT TAG GTG ACA TAG.

location/Qualifiers
1. 853
/organism="Nicotiana benthamiana"
/mol_type="mRNA"
/db_xref="taxon:4100"
/clone="NEMD912"
/tissue_type="abiotic and biotic stress-treated leaves,
callus tissue and root tissue"
/lab_host="DH10B-Tona"
/clone.lib="Nicotiana benthamiana mixed tissue cDNA
library, normalized, full-length"
/note="Vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI;
supplier: RNA was isolated from Nicotiana benthamiana
tissues that include callus, roots from liquid culture
grown plants, heat-stressed leaves (38 C, 3 hr and 6 hr),
cold-stressed leaves (5 C 3 hr, 6hr), and pathogen
challenged leaves (Pseudomonas syringae pv tomato 12 hr;
Pseudomonas syringae pv campestris 12 hr, 18hr;
Xanthomonas campestris pv vesicatoria 18hr). RNA was isolated from
these tissues and pooled in approximately equal molar
amounts."

ORIGIN
Query Match 10.7%; Score 34.6; DB 14; Length 853;
Best Local Similarity 47.5%; Pred. No. 25;
Matches 103; Conservative 0; Mismatches 114; Indels 0; Gaps 0;

Fukuda, S., Hashizume, M., Hayashida, K., Hirozane, T., Hori, F., Imocani, K., Ishii, Y., Itoh, M., Kagawa, I., Kawai, J., Kojima, Y., Kondo, S., Kono, H., Koya, S., Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Saito, R., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M., Takeda, Y., Waki, K., Watanabe, A., Muramatsu, M., and Hayashizaki, Y. Direct Substitution

Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome 12, 673-677 (2001)
Normalization and subtraction of cap-triapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multichannel sequencer. Genome Res. 10 (11), 1757-1771 (2000)

Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Tissues were provided by Dr. John Todd (Dept. of Medical Genetics Wellcome Trust Centre for Molecular Mechanisms in Disease Wellcome Trust/MRC Building Addenbrookes Hospital Cambridge) whose assistance we gratefully acknowledge.
Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

FEATURES

source

Location/Qualifiers
1..728
/organism="Mus musculus"
/mol_type="mRNA"
/strain="NOD"
/db_xref="taxon:10090"
/clone="F830034E14"
/tissue_type="activated spleen"
/clone_lib="RIKEN full-length enriched, activated spleen"

ORIGIN

Query Match 10.7%; Score 34.8; DB 13; Length 728;
Best Local Similarity 55.9%; Pred. No. 20;
Matches 66; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

QY 188 CGATGATCTGAGAGCTTAACCTTGTCTCATTCGCCGTGGATTACAGTGTGGCC 247

DB 244 CGAGCTGAGCTGATCTTTAATCCCTGGGTGTAAGCTGCTGTGTAAGAGAGCTGTGC 303

QY 248 ATCTTTGGCGGTGCGACGACCTTGTGAGCGTGGCGCTGGACCATCGTTTGT 305

DB 304 AACATCTGTGTGCGACCCCTCTTCTTGTGCGTGTGTGTAAGGCGACGCTTCTGT 361

RESULT 9

LOCUS

BI144284 808 bp mRNA linear EST 05-JUL-2001

DEFINITION 60290725F1 NCI CGAP_Kid14 Mus musculus cDNA clone IMAGE:5064373

ACCESSION BI144284

VERSION BI144284.1 GI:14604285

KEYWORDS

SOURCE

ORGANISM

Mus musculus (house mouse)

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Unpublished (1999)
National Institutes of Health, Mammalian Gene Collection (MGC)
NIH-MGC <http://mgc.nci.nih.gov/>
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
<http://image.lnl.gov>
Plate: L1M1173 row: m column: 14
High quality sequence start: 6
High quality sequence stop: 793.

FEATURES

source

Location/Qualifiers
1..808
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:5064373"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NCI CGAP Kid14"
/note="Organ: kidney; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.75 kb. Constructed by Life Technologies. Note: this is a NCI CGAP Library." "

ORIGIN

Query Match 10.7%; Score 34.8; DB 12; Length 808;
Best Local Similarity 53.7%; Pred. No. 21;
Matches 72; Conservative 0; Mismatches 62; Indels 0; Gaps 0;

QY 133 TGGATGCCAGCAGGATCTTGCCATTGTCAGCCGATCAAGTTTCGAGCATGCGATA 192

DB 523 TGGAGGCTTCATGTTCCAACTTAATGTGAGCGACAGCTGGAGCATCAACAGACATG 582

QY 193 GATCTGAAGACTTAACCTTGTCTCATTCGCCGTGGCATTAAGTGTGGCCACTT 252

DB 583 GTTCAGAGATGAGAGAGCTTGAGGTTCAAAAGCTTGTGTGCTACTGTAGATGCCCTGTA 642

QY 253 CTGGCGGTGCGACG 266

DB 643 CTGGCATATAGCG 656

RESULT 10

LOCUS

BX387771 1201 bp mRNA linear EST 08-MAY-2003

DEFINITION BX387771 Homo sapiens PLACENRA COT 25-NORMALIZED Homo sapiens cDNA clone CS01052YB18 5-PRIME. mRNA sequence.

ACCESSION BX387771 GI:30460721

KEYWORDS

SOURCE

ORGANISM

Homo sapiens (human)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 1201)

Full-length cDNA libraries and normalization

COMMENT

Genoscope - Centre National de Sequencage
BP 191 91066 Evry cedex - France
Email: segreg@genoscope.cns.fr, Web: www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 4682.r For more information about this cluster, see <http://www.genoscope.cns.fr/>
cgi-bin/cviewer.cgi?seq=CS01052DA090P1cluster=4682.r. Contact: Feng Liang Email: fliang@life-tech.com URL: <http://fulllength.invitrogen.com/> Invitrogen Corporation 1600 Faraday Avenue Genoscope sequence ID: CS01052DA090P1.

FEATURES

source

Location/Qualifiers
1..1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"

/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBTA530007"
/note="Vector: PCR4-TOPO; Site 1: EcoRI; 0.6-1.0 kb high
Cor selected genomic DNA library"

Query Match 10.9%; Score 35.2; DB 28; Length 910;
Best Local Similarity 57.1%; Pred. No. 17;
Matches 64; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

Db 192 AGATCTGAAGACTTACCTTTGGTCTCATTCGCTTGCATTAACGTGTGGCGCATCT 251
272 AAATCTACGCGCTCAATCTCTTCCCAATCCGACATCGCGACGTCGCGCGCC 331

Qy 252 TCTTGGCGGTCGACGACCTTGTGAGCGTGGCGCTGACCATGCTTTT 303
Db 332 TGTGTGGCGGCGATCTCGCTGGCGACGACCGGATGCTGCGCGCTTT 383

RESULT 7 1101 bp DNA linear GSS 04-JUN-1999
CNS00BD8 Drosophila melanogaster genome survey sequence TERT end of BAC #
LOCUS BACR23H11 of RPI1-98 library from Drosophila melanogaster (fruit
DEFINITION fly), genomic survey sequence.
ACCESSION AL056828.1 GI:4937496
VERSION AL056828
KEYWORDS GSS.
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 1101)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 Evry cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)

COMMENT Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see <http://www.fruitfly.org> The BDGP Drosophila
melanogaster BAC library was prepared by Kazuo Ooegawa and
Aaron Mamoser in Peter de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPI1-98 and was constructed by partial
EcoRI digestion of Drosophila DNA provided by the BDGP from the
isogenic strain Y2; cn bw sp, the same strain used for the BDGP's
P1 and EST libraries. A more detailed description of the library
and how to order individual BAC clones, the entire library, or
filters for hybridization from the BACPAC Resource Center can be
found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES
Location/Qualifiers
1..1101

Source
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACR23H11"
/clone_lib="RPI1-98"
/note="end : TERT"

ORIGIN
Query Match 10.9%; Score 35.2; DB 29; Length 1101;
Best Local Similarity 24.4%; Pred. No. 18;
Matches 62; Conservative 81; Mismatches 111; Indels 0; Gaps 0;

Qy 29 TTGTTGTCAGTATGTCAGTCACTTTTGGCGTCCGGCGGCTCCGTTCTTATCC 88
Db 847 TTGSGKTGTGSGAASDGTGTTATTTTTSKTSBTSKMAATKWSASTTTT 906

Qy 89 TTAAGCCCTTACGTATCAATTTGTGGGCAAAATGCGATGTGATGCCACAGAA 148
Db 907 WSGASGTSITBTKXSGTSTGTTTWSGSAVAASGTTTWTSTWSTBKABBA 966
Qy 149 TCTTGGCCATTTTGGACGCATCAACGTTTCGACCAATGCGATGATCTGAAAGACTCTAA 208
Db 967 ATSAHTABSDPTAABSKTSBSBSSDDBGTRASGSAATTTKTTTITTTTWW 1026
Qy 209 CTTTGTCTCATTCGCCCTTGCATTAACGTGTGGCGCATCTTTTGGCGTGCAGCA 268
Db 1027 TATTTNSBSBWBATTTTDTAAATWAATVSTNSNSNNNTTTBSSRAATPAABVSA 1086
Qy 269 CTTTGTACGCTT 282
Db 1087 SGTBGTBRBSVATB 1100

RESULT 8 728 bp mRNA linear EST 17-DEC-2002
BY765711
LOCUS BY765711 RIKEN full-length enriched, activated spleen Mus musculus
DEFINITION cDNA clone F830034E14 3', mRNA sequence.
ACCESSION BY765711
VERSION BY765711.1 GI:27203910
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 728)
AUTHORS Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S.,
Nikaido, I., Oato, N., Saito, R., Suzuki, H., Yamana, A.,
Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A.,
Schonbach, C., Gojobori, T., Baldarelli, R., Hill, D.P., Bult, C.,
Hume, D.A., Quackenbush, J., Schmitt, L.M., Kanapin, A., Matsuda, H.,
Batalov, S., Beisel, K.W., Blake, J.U.A., Bradt, D., Brasic, V.,
Chochia, C., Corbani, L.E., Cousins, S., Daille, E., Dragani, T.A.,
Fletcher, C.F., Forrest, A., Frazer, K.S., Gaasterland, T.,
Gariboldi, S., Giesi, C., Godzik, A., Gough, J., Gilmord, S.,
Gustinch, S., Hirokawa, N., Jackson, I.J., Jarvis, E.D., Kanai, A.,
Kawaji, H., Kawasawa, Y., Kedzierski, R.M., King, B.L., Kongave, A.,
Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglocz, D.R.,
Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T.,
Numata, K., Okido, T., Pavan, W.J., Petrea, G., Pesole, G.,
Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D., Ramchandran, S.,
Ravasi, T., Reed, J.C., Reed, D.J., Reid, J., Ring, B.Z., Ringwald, M.,
Sandelin, A., Schneider, C., Semple, C.A., Setou, M., Shimada, K.,
Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale, R.D., Tomita, M.,
Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y., Warande, Y.,
Wells, C., Wilming, L.G., Wyshaw-Boris, A., Yanaigawa, M., Yang, I.,
Yang, L., Yuan, Z., Zavalan, M., Zhu, Y., Zimmer, A., Carninci, P.,
Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura, M.,
Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K.,
Arikawa, T., Fukuda, S., Hara, A., Hashizume, W., Imocani, K., Ishii, Y.,
Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K.,
Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E.S.,
Rogers, J., Birney, E. and Hayashizaki, Y.

TITLE Analysis of the mouse transcriptome based on functional annotation
JOURNAL Nature 420, 563-573 (2002)
MEDLINE 22354683
PUBMED 12468851
COMMENT Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suenho-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@ac.riken.go.jp,
URL: <http://genome.gsc.riken.go.jp/>
Adachi, J., Aizawa, K., Akimura, T., Arikawa, T., Carninci, P.,

LOCUS CF338060 475 bp mRNA linear EST 18-AUG-2003
DEFINITION UMYT-08-M05.b1 AtUMT-overexpressing transgenic rice plasmid cDNA library (UMT) *Oryza sativa* cDNA clone UMT-08-M05, mRNA sequence.
ACCESSION CF338060
VERSION
KEYWORDS CF338060.1 GI:33824491
SOURCE EST.
ORGANISM *Oryza sativa*
Oryza sativa
Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehharctoidae; Oryzaceae; Oryza.
REFERENCE 1 (bases 1 to 475)
Kim,U.S., Jun,K.M., Cheong,P.J., Kim,M.J., Lee,T.H., Shin,Y.C., Song,S.I., Kim,J.K., Kim,Y.-K. and Nahm,B.H.
Large-scale Sequencing Analysis of Rice ESTs
Unpublished (2003)
COMMENT Contact: Nahm B.H.
Genomics and Genetics Institute, Greengene Biotech Inc., Division of Bioscience and Bioinformatics, Yonsei University
Yongin, Kyeonggi, Korea
Tel: 82 31 330 6193
Fax: 82 31 321 6355
Email: bhnahm@gbio.com, bhnahm@bio.myongji.ac.kr.
FEATURES
Source 1..475
Location/Qualifiers

```

FEATURES
    source
        location/qualifiers
            1..475
                /organism="Oryza sativa"
                /mol_type="mRNA"
                /cultivar="Nackdong"
                /db_xref="taxon:4530"
                /clone="JMT-08-M05"
                /tissue_type="leaf"
                /dev_stage="14 days after germination"
                /lab_host="E.coli DH10B"
                /clone_lib="AluJMT-overexpressing transgenic rice plasmid
cdna library (JMT)"
                /note="vector: pCR4-TOP0; Site 1: EcoRI; Oligo-capped mRNA
was reverse transcribed and then used for PCR. mRNA was
pared from Arabidopsis usamosin Carboxyl
methyltransferase overexpression line."

```

	Query Match	10.9%	Score 35.2;	DB 14;	Length 475;
	Best Local Similarity	50.6%;	Pred. No. 13;		
	Matches	85;	Conservative	0;	Mismatches 83; Indels 0; Gaps 0;
Qy	51	CATTACTTTTGCGCTCCGGGCGGTTCGGTCTTAATCCTTAAGCCCCACGTGAATACA	110		
Db	255	CCTTCTTTTGGGCTCCCGGGGCTTCTGCTAAGCCTTTTAAGCGGCCCTTGCGGGCGC	196		
Qy	111	ATTGTGGGCMAAATGAGATGTGATGCCAGCAAGAAATCCTTGCAATTTGACCGCATC	170		
Db	195	CGGGGTGTCCGCCCTCATGAGGTTGGCAGCGGCTCTTGCAAGATCGCGAGGTTT	136		
Qy	171	AACGTTTGCAGCAATGCGATAGATCTTAAGAATCTTAACCTTTGGTCT	218		
Db	135	CACGCGCTTCAGAGAAAGCAATTCCTTGATGATCTTCTGCTGTCTT	88		
RESULT 5					
BZ720836					
LOCUS	BZ720836	786 bp	DNA	linear	GSS 24-FEB-2003
DEFINITION	BZ720836	PUCFWM5TD ZM_0.6_1.0_KB Zee mays genomic clone ZMMB1A137A09,			
ACCSSION	BZ720836	genomic survey sequence.			
VERSION	BZ720836.1	GI:28512409			

KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Zea mays
Spermatophytes; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 786)
Whitelaw, C.A., Quackenbush, J., Van Aken, S., Uttreich, T.,

```

TITLE
JOURNAL
COMMENT
Renick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
Benetzen,J.
Maize Genomics Consortium
Unpublished (2003)
Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@cigr.org
Seq primer: TF
Class: sheared ends.

FEATURES
    source
        location/Qualifiers
            1..786
                /organism="Zea mays"
                /mol_type="genomic DNA"
                /strain="B73"
                /db_xref="taxon:4577"
                /clone="ZMMER137A09"
                /clone_1db="ZM_0.6_1.0_KB"
                /note="Vector: PCR4-TOPO; Site_1: EcoRI; 0.6-1.0 kb high
Cor selected genomic DNA library"

ORIGIN

Query Match      10.9%; Score 35.2; DB 28; Length 786;
Best Local Similarity 53.7%; Pred.No.16;
Matches 73; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

```

	Query Match	10.9%;	Score 35.2;	DB 28;	Length 786;
	Beet Local Similarity	53.7%;	Pred. No. 16;		
	Matches 73;	Conservative	0;	Mismatches 63;	Indels 0;
				Gaps	0;
Qy	189 GATGATCTGAGACCTTAACCTTGGCTTCATTGCGCGTGCATTCACAGTGGGGCA	248			
Db	480 GAAAGCATCAAAAAATTGTGTCTAATCTCAAGTGGTGGCCAACTGTGTGGCTTA	539			
Qy	249 TCTTCTTGGCGGTGCAGCACTTTGTTAGCGTTGGCGCTGGCCACATCGTTTGTGG	308			
Db	540 TCATTAGTGGTGATCTTGAATGATTATCATTTGAGCCCAACAATCTTGTGAGTGG	599			
Qy	309 ACTGGTGAATCTTTTC	324			
Db	600 TATGTTGCACATGTGC	615			

```

RESULT 6
CC362201
LOCUS
DEFINITION
CC362201
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Zea mays
Zea mays
Eukaryotes; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoidae; Andropogoneae; Zea.
1 (bases 1 to 910)
Whitefaw,C.A., Quackenbush,T., Van Aken,S., Uterback,T.,
Reinick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
Bennerzen,J.
Maize Genomics Consortium
Unpublished (2003)
Other_GSSs: PHTK887B
Contact: Cathy Whitefaw
TIR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitefaw@tigr.org
Seq primer: TP
Class: sheared ends.
Location/Qualifiers
1..910
/organism="Zea mays"
/mol type="genomic DNA"

```


Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of *Drosophila* DNA provided by the BDGP from the isogenic strain y2; cn bw sp, the same strain used for the BDGP's p1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES

Location/Qualifiers
1..995
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACR12G04"
/clone_id="RPCI-98"
/note="end : TET3"

ORIGIN

Query Match 11.7%; Score 37.8; DB 29; Length 995;
Best Local Similarity 34.0%; Pred. No. 2.9;
Matches 54; Conservative 41; Mismatches 64; Indels 0; Gaps 0;

QY 164 CCGCATCAACGTTTGGAGCAATCGATGATCTGAAGCTTAACTTGGTCTGATG 223
DB 872 CAGTMTTACGACCGGSGTGTGMAAAGSASASATGVRSTCGCTTTMTTATAT 813
QY 224 CCGTTGGATTACAGTGTGGCATCTCTTGGCGGTCGACGACCTTGTGAGCGTG 283
DB 812 CSKTSBTTWGKASCMTTSSGSMASWTGTGKKGAGSAGAAAGGCGCAKKS 753
QY 284 GCGCTGCACCATGCTTTTGTGACTGTGAATCTT 322
DB 752 SSGTGTGKXTSARKKTTGTGTGSGTGATCTTT 714

RESULT 2 1353 bp DNA linear GSS 17-DEC-2002
B2557685
LOCUS pasc1-60.663, g1 pasc1-60 Pseudomonas aeruginosa genomic clone
DEFINITION pasc1-60.663, genomic survey sequence.
ACCESSION B2557685
VERSION B2557685.1 GI:27170831
KEYWORDS GSS.
SOURCE Pseudomonas aeruginosa
ORGANISM Pseudomonas aeruginosa
Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales; Pseudomonadaceae; Pseudomonas.
REFERENCE 1 (bases 1 to 1353)
Spencer, D.H., Raymond, C.K., Smith, E.E., Sims, E.E., Hastings, M., Burns, J.L., Kaul, R. and Olsen, M.V. Whole-Genome-Sequence Variation among multiple isolates of *Pseudomonas aeruginosa* library J. Bacteriol. (2002) In press
CONTACT: Chris K. Raymond
Genome Center
University of Washington
Box 353145, Seattle, WA 98105-2145, USA
Tel: 2062216954
Fax: 2066857244
Email: ckraymond@u.washington.edu
Class: shotgun.

FEATURES

Location/Qualifiers
1..1353
/organism="Pseudomonas aeruginosa"
/mol_type="genomic DNA"
/strain="1-60"
/db_xref="taxon:287"
/clone="pasc1-60.663"
/clone_id="pasc1-60"
/note="Clinical isolate 1-60 Whole genomic shotgun library."

ORIGIN

library."

Query Match 11.0%; Score 35.8; DB 28; Length 1353;
Best Local Similarity 57.7%; Pred. No. 13;
Matches 64; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

QY 204 TCTAACCTTGTGCTCATTTGCGGATTCAGATGAGTGGCGCATCTTGGCGGTG 263
DB 86 TATAGCCCTTCTCTCGGCGGTTGAAGTGCCTGGCGCGCTTATCTGCTGATCC 145
QY 264 AGCACCCTTGTGAGCGTGGCGCTGCACCATGCTTTTGTGCACTGG 314
DB 146 GCGCGAGGGGCTTCTCATCGGAGTGTGTCGCGCTTGTGCGGACAGG 196

RESULT 3

CP302597/c 471 bp mRNA linear EST 15-AUG-2003
LOCUS 7LEAF--08-E15.b1 Rice leaf plasmid cDNA library II (7LEAF) Oryza
DEFINITION sativa cDNA clone 7LEAF--08-E15, mRNA sequence.
ACCESSION CP302597
VERSION CP302597.1 GI:33674358
KEYWORDS EST.
SOURCE Oryza sativa
ORGANISM Oryza sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.
REFERENCE 1 (bases 1 to 471)
Kim, J.S., Jun, K.M., Cheong, P.J., Kim, M.J., Lee, T.H., Shin, Y.C., Song, S.I., Kim, J.K., Kim, Y.-K. and Nam, B.H. Large-scale Sequencing Analysis of Rice ESTs Unpublished (2003)
CONTACT: Nam, B.H.
Genomics and Genetics Institute, GreenGene Biotech Inc., Division of Bioscience and Bioinformatics, Myongji University
Yongin, Kyonggi, Korea
Tel: 82 31 330 6193
Fax: 82 31 321 6355
Email: bhnam@bio.com, bhnam@bio.myongji.ac.kr.

FEATURES

Location/Qualifiers
1..471
/organism="Oryza sativa"
/mol_type="mRNA"
/cultivar="Nackdong"
/db_xref="taxon:4530"
/clone="7LEAF--08-E15"
/issue_type="leaf"
/dev_stage="7 days after germination"
/lab_host="E.coli DH10B"
/clone_id="Rice leaf plasmid cDNA library II (7LEAF)"
/note="Vector: pCR4-TOPO; Site 1: EcoRI; mRNA was capped with oligoribonucleotides and then used as templates for RT-PCR."

ORIGIN

Query Match 10.9%; Score 35.2; DB 14; Length 471;
Best Local Similarity 50.6%; Pred. No. 13;
Matches 85; Conservative 0; Mismatches 83; Indels 0; Gaps 0;

QY 51 CATTACTTTTGGGCTCCGGGCGGCTTCCGTTCTTAATCCTTAGCCCTACGAGATCA 110
DB 224 CTTCTCTTGGGCTCCGGGCGGCTTCTGCTGACGCTTCTAGCGCCCTTGGGCGG 165
QY 111 ATTGTGGGCAAAATGCGATGTGATGCGACAGAGATCCTTGCCATTTTACCGCATC 170
DB 164 CGGCGTTTGGCGGCTGTACTAGGCTGCGACCGGCTCTTGCGAAGTCCGCGAGTTT 105
QY 171 AACGTTTGGCAGATGATGATCTGAAGACTTGAAGCTTGTGCT 218
DB 104 CACGGCTTCAGAGAGCCATCTTGTAGTCTTCTGCTGTGCT 57

RESULT 4

CF338060/c

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 17, 2004, 06:48:18 ; Search time 1049.18 Seconds

(without alignments)
9221.792 Million cell updates/sec

Title: US-10-608-504-6_COPY_853_1176

Perfect score: 324
Sequence: 1 atgacacacgattcttcctg.....ttgacgtgtaacctttc 324

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST.*
1: em_estda:*
2: em_estnum:*
3: em_estlin:*
4: em_estm:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_estc:*
9: gb_estl:*
10: gb_est2:*
11: gb_est3:*
12: gb_est4:*
13: gb_est5:*
14: gb_est6:*
15: em_estom:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_man:*
23: em_gss_mus:*
24: em_gss_pio:*
25: em_gss_rtd:*
26: em_gss_pdg:*
27: em_gss_vrl:*
28: gb_gssl:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	37.8	11.7	995	29	CNS005NG
C 2	35.8	11.0	1353	28	BZ557685 pasc1-60
C 3	35.2	10.9	471	14	CF302597 7LEAF--08
C 4	35.2	10.9	475	14	CF338060 JMT--08-M

Result No.	Score	Query Match	Length	ID	Description
C 5	35.2	10.9	786	28	BZ720836
C 6	35.2	10.9	910	28	CC362201
C 7	35.2	10.9	1101	29	CNS005NG
C 8	34.8	10.7	728	13	BY765711
C 9	34.8	10.7	808	13	BY765711
C 10	34.8	10.7	1201	13	BY765711
C 11	34.6	10.7	413	13	CB239541
C 12	34.6	10.7	853	14	CK297477
C 13	34.6	10.7	890	14	CK286648
C 14	34.6	10.7	922	14	CK295854
C 15	34.6	10.7	936	14	CK295474
C 16	34.6	10.7	951	14	CK292130
C 17	34.4	10.6	447	29	CNS02NHL
C 18	34.4	10.6	1201	13	BY356376
C 19	34.2	10.6	1062	28	BZ557686
C 20	34.2	10.5	423	28	B44402
C 21	33.8	10.4	1088	29	CNS046RC
C 22	33.6	10.4	833	13	BY588308
C 23	33.4	10.3	580	12	BM309296
C 24	33.2	10.2	478	14	H65669
C 25	33.2	10.2	586	14	CD091362
C 26	33.2	10.2	859	10	BF341287
C 27	33.2	10.2	1044	28	BZ553873
C 28	33.2	10.2	569	12	BM100089
C 29	33.2	10.2	720	29	CG081864
C 30	33.2	10.2	761	29	CG084816
C 31	33.2	10.2	796	29	CG084833
C 32	33.2	10.2	818	28	CC005623
C 33	33.2	10.2	958	29	CC084842
C 34	33.8	10.1	390	10	AM456947
C 35	33.8	10.1	399	10	AM491785
C 36	32.8	10.1	416	10	AM457742
C 37	32.8	10.1	463	10	AM489067
C 38	32.8	10.1	469	10	BF460793
C 39	32.8	10.1	486	12	BI135202
C 40	32.8	10.1	496	12	BI134847
C 41	32.8	10.1	536	12	BI901653
C 42	32.8	10.1	565	12	BI901912
C 43	32.8	10.1	602	10	BF207637
C 44	32.8	10.1	638	14	CF582833
C 45	32.8	10.1	3361	11	AK083916

ALIGNMENTS

RESULT 1
CNS005NG/c
LOCUS
DEFINITION
Drosophila melanogaster genome survey sequence TERT end of BAC #
BACR12604 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.
ACCESSION
AL060428.1 GI:4943359
VERSION
KEYWORDS
SOURCE
ORGANISM
Drosophila melanogaster (fruit fly)
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE
Genoscope.
Direct Submission
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see <http://www.fruitfly.org> The BDGP Drosophila melanogaster BAC library was prepared by Kazuo Oosawa and Aaron Mammoser in Pieter de Jong's laboratory in the Department of

QY 178 CGGAGCAATGCGATAGATCTGAAAGACTCTTAACCTTTGCTCATGCGCTGGGATTACA 237
DB 15721 GGGTCCACTCTGCGGGGAAGAGGCGCTTGAAAGACAGACACCATCTGCTGCGGTGTGA 15780
QY 238 GTGATGCGGCGATCTTCTTGGCG 259
DB 15781 GTGTACTCTATCTTCTGCTG 15802

RESULT 14
US-10-796-280-12378
; Sequence 12378, Application US/10796280
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: C1001510
; CURRENT APPLICATION NUMBER: US/10/796,280
; CURRENT FILING DATE: 2004-03-10
; NUMBER OF SEQ. ID NOS: 68533
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ. ID NO 12378
; LENGTH: 234309
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(234309)
; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1-
US-10-796-280-12378

Query Match 8.8%; Score 28.4; DB 6; Length 234309;
Best Local Similarity 52.5%; Pred. No. 81;
Matches 62; Conservative 0; Mismatches 56; Indels 0; Gaps 0;

QY 206 TAACTTTGCTCATGATCCGCTTGCATTTACATGCTGGCGCATCTTTTGGCGGTGAC 265
DB 62038 TCACCTTAGGCGCTTTAAGCCCTGAAGAGCGTGGCGCTTACATCTGCTTTTATTTTCA 62097
QY 266 GCACCTTGTAGCGGTGGCGCTGACCATCGTTTGTGACCTGATGATCTTTT 323
DB 62098 GTCCTTGTGTAGTGTGCTAAGGCGCTTTCGCTCTGTGTAGAGAGGGGTTT 62155

RESULT 15
US-10-767-701-27384/C
; Sequence 27384, Application US/10767701
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ. ID NOS: 63128
; SEQ. ID NO 27384
; LENGTH: 490
; TYPE: DNA
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: 6858297
US-10-767-701-27384

Query Match 8.7%; Score 28.2; DB 6; Length 490;
Best Local Similarity 51.2%; Pred. No. 5.8;
Matches 66; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

QY 192 AGATCTGAAGACTTAACTTGGTCTCATTTGGCGGTGACATTACAGTGTGGCGCATCT 251
DB 282 AGGTGTACAGGTGTAGCACTTGTGTAGCTGCCAGAGGACGGAGTGTAGCCAGCC 223

QY 252 TCTTGGCGGTGACGACCACTTGTGAGGCTTGGCGGTGACCACTGTTTGTGACT 311
DB 222 AGTTGGCGGTGACGAGGTGACGACTGCTGGCGGTGACCACTGAGACTTGTGACT 163
QY 312 GGTGAATCT 320
DB 162 GGGCAACT 154

Search completed: April 17, 2004, 20:13:04
Job time: 39.4407 secs

QY 143 CAGGATCTTGGCATTGTAACCGCATCAACGTTTCCAGCAATGCATGATCTGAGA 202
 DB 284 CGGGTATTTGGCGGAGATATACCGACGACTATTCAACCGCATAGGTTTGGCCGACGG 225
 QY 203 CTCGACCTTGGCTCATTTGCGGTGCGATTACAGTGTGGCGCATCTTCTTGGCGGTC 262
 DB 224 CACTTAACCGTAGCGGATTTGGCTTCCAGCACTGATTTGGCTTCAACGTGGCGATG 165
 QY 263 GACGACCTTGTG 276
 DB 164 TTCTGATCATGATG 151

RESULT 10
 US-10-108-605A-248/c
 ; Sequence 248, Application US/10108605A
 ; GENERAL INFORMATION:
 ; APPLICANT: Broadus, Julie
 ; APPLICANT: Stam, Lynn
 ; APPLICANT: Bachmann, Jane
 ; APPLICANT: Kamdar, Kim
 ; TITLE OF INVENTION: NUCLEIC ACID SEQUENCES FROM DROSOPHILA MELANOGASTER THAT ENCODE
 ; TITLE OF INVENTION: PROTEINS ESSENTIAL FOR LARVAL VIABILITY AND USES THEREOF
 ; FILE REFERENCE: 31133B
 ; CURRENT APPLICATION NUMBER: US/10/108, 605A
 ; PRIOR FILING DATE: 2002-03-27
 ; PRIOR APPLICATION NUMBER: US 09/761,142
 ; PRIOR FILING DATE: 2001-01-16
 ; PRIOR APPLICATION NUMBER: US 60/176,418
 ; NUMBER OF SEQ ID NOS: 361
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 248
 ; LENGTH: 8577
 ; TYPE: DNA
 ; ORGANISM: Drosophila melanogaster
 US-10-108-605A-248

Query Match 8.8%; Score 28.4; DB 6; Length 8577;
 Best Local Similarity 60.3%; Pred. No. 18;
 Matches 47; Conservative 0; Mismatches 31; Indels 0; Gaps 0;
 QY 209 CTTTGTCTCATTTCCGTTGCGATTACAGTGTGGCGCATCTTCTTGGCGGTGACGA 268
 DB 83 CCGTGTGTGTGTGCGCCGACGCTACTGTTGTGATGCTTTGGCGAGGACGAC 24
 QY 269 CTTGTGAGCGTTGCG 286
 DB 23 CCGCTTCAACGCGGTG 6

RESULT 11
 US-10-796-280-12417/c
 ; Sequence 12417, Application US/10796280
 ; GENERAL INFORMATION:
 ; APPLICANT: CARGILL, Michele et al.
 ; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
 ; TITLE OF INVENTION: STENOSIS, METHODS OF DETECTION AND USES THEREOF
 ; FILE REFERENCE: CL001510
 ; CURRENT APPLICATION NUMBER: US/10/796,280
 ; PRIOR FILING DATE: 2004-03-10
 ; NUMBER OF SEQ ID NOS: 68533
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 12417
 ; LENGTH: 20618
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-10-796-280-12417

Query Match 8.8%; Score 28.4; DB 6; Length 20618;
 Best Local Similarity 51.6%; Pred. No. 27;
 Matches 65; Conservative 0; Mismatches 61; Indels 0; Gaps 0;

QY 65 TCCGGCGGTTCCGTTCTTAATCCTTAAGCCCTAGCGATACAAATTTGGGCAAA 124
 DB 8143 TCCATCATTTCCACTGTATTAACCTTTGGCTCAAGCAGATACCTGTTGGATTAA 8084
 QY 125 TGGCATGTGATGACGACAGAGAAATCCTTGCCATTTTGACCGCATCAACGTTTGGACGA 184
 DB 8083 TGCTGTCTCACTGTTTCATTCCTGCGCCCTGCGCCCTGTTCCCAACTATTTCTACCA 8024
 QY 185 ATGCA 190
 DB 8023 AGGCA 8018

RESULT 12
 US-10-796-280-12418
 ; Sequence 12418, Application US/10796280
 ; GENERAL INFORMATION:
 ; APPLICANT: CARGILL, Michele et al.
 ; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
 ; TITLE OF INVENTION: STENOSIS, METHODS OF DETECTION AND USES THEREOF
 ; FILE REFERENCE: CL001510
 ; CURRENT APPLICATION NUMBER: US/10/796,280
 ; PRIOR FILING DATE: 2004-03-10
 ; NUMBER OF SEQ ID NOS: 68533
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 12418
 ; LENGTH: 46075
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-10-796-280-12418

Query Match 8.8%; Score 28.4; DB 6; Length 46075;
 Best Local Similarity 51.6%; Pred. No. 39;
 Matches 65; Conservative 0; Mismatches 61; Indels 0; Gaps 0;
 QY 65 TCCGGCGGTTCCGTTCTTAATCCTTAAGCCCTAGCGATACAAATTTGGGCAAA 124
 DB 3888 TCCATCATTTCCACTGTATTAACCTTTGGCTCAAGCAGATACCTGTTGGATTAA 3947
 QY 125 TGGCATGTGATGACGACAGAGAAATCCTTGCCATTTTGACCGCATCAACGTTTGGACGA 184
 DB 3948 TGCTGTCTCACTGTTTCATTCCTGCGCCCTGCGCCCTGTTCCCAACTATTTCTACCA 4007
 QY 185 ATGCA 190
 DB 4008 AGGCA 4013

RESULT 13
 US-10-796-280-12452
 ; Sequence 12452, Application US/10796280
 ; GENERAL INFORMATION:
 ; APPLICANT: CARGILL, Michele et al.
 ; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
 ; TITLE OF INVENTION: STENOSIS, METHODS OF DETECTION AND USES THEREOF
 ; FILE REFERENCE: CL001510
 ; CURRENT APPLICATION NUMBER: US/10/796,280
 ; PRIOR FILING DATE: 2004-03-10
 ; NUMBER OF SEQ ID NOS: 68533
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 12452
 ; LENGTH: 69359
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-10-796-280-12452

Query Match 8.8%; Score 28.4; DB 6; Length 69359;
 Best Local Similarity 50.0%; Pred. No. 47;
 Matches 71; Conservative 0; Mismatches 71; Indels 0; Gaps 0;
 QY 118 GGCATAATGGCATGTGATGACGAGAAATCCTTGCCATTTTGACCGCATCAACGTTT 177
 DB 15661 GGCATAATGGCATGTGATGACGAGAAATGAGCATGTTTACCTCAGAGTCTGTGCCCCCTGTGCATT 15720

Query Match	8.8%;	Score 28.4;	DB 1;	Length 500;
Best Local Similarity	50.7%;	Pred. No. 4.9;		
Matches	68;	Conservative	0;	Mismatches 66;
			Indels	0;
			Gaps	0;

US-10-789-400-36
; Sequence 36, Application US/10789400
; GENERAL INFORMATION:
; APPLICANT: THE GOVERNMENT OF THE UNITED STATES OF AMERICA AS
; APPLICANT: REPRESENTED BY THE SECRETARY OF THE DEPARTMENT OF HEALTH AND
; APPLICANT: HUMAN SERVICES
; APPLICANT: COLLINS, Peter L.
; APPLICANT: Biacchesi, Stephane
; APPLICANT: Buchholz, Ursula
; APPLICANT: Skidopoulos, Mario H.
; APPLICANT: Murphy, Brian R.
; TITLE OF INVENTION: RECOMBINANT HUMAN METAPNEUMOVIRUS AND ITS USE
; FILE REFERENCE: 4239-67783
; CURRENT APPLICATION NUMBER: US/10/789, 400
; PRIOR FILING DATE: 2004-02-27
; PRIOR APPLICATION NUMBER: US 60/451,119
; PRIOR FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/478,667
; PRIOR FILING DATE: 2003-06-13
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 36
; LENGTH: 13350
; TYPE: DNA
; ORGANISM: human metapneumovirus
US-10-789-400-36

Query Match 9.4%; Score 30.4; DB 6; Length 13350;
Best Local Similarity 49.4%; Pred. No. 4.5;
Matches 79; Conservative 0; Mismatches 81; Indels 0; Gaps 0;

QY 14 TCTCCGTATCTCTCTGTTGTCGAGTATGTCAGTCATCTTTGGCCCTCGGGCGG 73
DB 1687 TCTCAGATATGAAGAAGAAAGATGACAGATCTTCACTTGAAGAAAGAGATA 1746
QY 74 TTCGTTCTTATCTTAAAGCCCTACGATGATGATGATGATGATGATGATGATGATGAT 133
DB 1747 CTTATCTTATGAGATTCAGCCGATGATGATGATGATGATGATGATGATGATGAT 1806
QY 134 GGAATGCCAGAGAAATCTTGGCATTTTACCGCATAC 173
DB 1807 TAGGGCTATTAAGAACTCAACATTTGCTACAGCAGAC 1846

RESULT 3
US-10-767-471-10749/c
; Sequence 10749, Application US/10767471
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: RHEUMATOID ARTHRITIS, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001505
; CURRENT APPLICATION NUMBER: US/10/767,471
; CURRENT FILING DATE: 2004-01-30
; NUMBER OF SEQ ID NOS: 50231
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10749
; LENGTH: 35895
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURES:
; NAME/KEY: misc_feature
; LOCATION: (1)...(35895)
; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1-
US-10-767-471-10749

Query Match 9.3%; Score 30.2; DB 6; Length 35895;
Best Local Similarity 52.8%; Pred. No. 8.2; 56; Indels 0; Gaps 0;
Matches 65; Conservative 0; Mismatches 56; Indels 0; Gaps 0;
QY 201 GACTCTAACCTTGTCTCATTCGCGTTCGATTAACAGTGTGGCGCATCTTCTGGCGG 260
DB 12662 GCCTTCACCCACCTCTCTCTGCTTGTCTTGGCAGACAGGCCATCTCTCTGCTG 12603

QY 261 TCGACGACCTTGTGAGCGTTGGCGCTGGACCACTGTTTGTGGAGTGAATCT 320
DB 12602 CCCACACATCTGCTTGGCCTCCTAGCTCCAGACAGAGATGTTTGAACACTAGACCT 12543
QY 321 TTT 323
DB 12542 TCT 12540

RESULT 4
US-10-767-701-5423/c
; Sequence 5423, Application US/10767701
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(5353)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 5423
; LENGTH: 731
; TYPE: DNA
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: SORBI-28WAY03-CLUS132149_1
US-10-767-701-5423

Query Match 8.9%; Score 28.8; DB 6; Length 731;
Best Local Similarity 53.6%; Pred. No. 4.3; 52; Indels 0; Gaps 0;
Matches 60; Conservative 0; Mismatches 52; Indels 0; Gaps 0;
QY 149 TCCITGCCATTTTGAACCGCATCAACGTTTGCAGCAATGATGATGATGATGATGATGAT 208
DB 240 TCTCGCCAGTTTGTCTCGGACACATCCACCATGGAATGGTTGACAGATTAACATG 181
QY 209 CCTTGGTCTCATTCCTGCGATTAACAGTGTGGCGCATCTTCTGGCGG 260
DB 180 CATTCATCATCTTGAAGTGTGAGTGTGAGCTGAACCAACTCTCGGCGTG 129

RESULT 5
US-10-767-471-10669/c
; Sequence 10669, Application US/10767471
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: RHEUMATOID ARTHRITIS, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001505
; CURRENT APPLICATION NUMBER: US/10/767,471
; CURRENT FILING DATE: 2004-01-30
; NUMBER OF SEQ ID NOS: 50231
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10669
; LENGTH: 158027
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURES:
; NAME/KEY: misc_feature
; LOCATION: (1)...(158027)
; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1-
US-10-767-471-10669

Query Match 8.9%; Score 28.8; DB 6; Length 158027;
Best Local Similarity 54.0%; Pred. No. 50; 44; Indels 0; Gaps 0;
Matches 54; Conservative 2; Mismatches 44; Indels 0; Gaps 0;
QY 48 AGTCATTAATCTTGGCGCTCGGCGGTTCCGTTCTTAATCTTAAGCCCTTAAGTGAATC 107
DB 48997 AGCTTTGAATCTTAACCAAGCTGCTATGATGATGATGATGATGATGATGATGATGAT 48938

Tue Apr 20 06:47:26 2004

us-10-608-504-6_copy_853_1176.rtf

Page 7

ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &
ADDRESSER: No. 6031150-ris
STREET: One Liberty Place 46th. Floor
CITY: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19103
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/476,900A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Beardell, Lori Y.
REGISTRATION NUMBER: 34,293
REFERENCE/DOCKET NUMBER: EYLR-0027
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-568-3100
TELEFAX: 215-568-3439
INFORMATION FOR SEO ID NO.: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 5110 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: nucleic acid
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-476-900A-4

	Query Match	Best Local Similarity	Score	DB	Length
	52	58.4%	29.8	3	5110
	Matches	Conservative	Pred. No. 3	Indels	Gaps
	0	0	37	0	0
Qy	118	GGCAAAATGGCGATGTGATCGACAGAAATCCTTGCCATTGTGACCGCATCAACGTTT	177		
Db	4500	GGCACTTGGCGATCTCAATGCCATCGATGATGATAGGCCCTTCAAAATATCCACCAATT	4441		
Qy	178	CGCAGCAATGCGATGATCTGAGACTCT	206		
Db	4440	CGAATAAAGCGAGAGAGAGAGGATTT	4412		

Search completed: April 19, 2004, 18:38:22
Job time : 28.2807 secs


```

MOLECULE TYPE: CDNA
HYPOTHEICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
FEATURE:
NAME/KEY: CDS
LOCATION: 37..4533
US-08-476-900A-27

Query Match      9.2%; Score 29.8; DB 3; Length 4635;
Best Local Similarity 58.4%; Pred. No.2.8;
Matches 52; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

Cy      118 GGCAAAATGGCGATGTGATGCCAGCAGGAATCCCTTGCCATTGTGACCGCATCAACGTTT 177
          |||||
Db      4025 GGCGACTGGGAGATGTCAATCCATTCGATGATGATAGCCCTTCAAACATATCCACCATT 3966
          |||||

Cy      178 CGCAGCAATGGCGATAGATCTGAAGACTCT 206
          |||||
Db      3965 CGGAAAAAGCGAGAGAGAAAGAGGATT 3937
          |||||

RESULT 12
US-08-488-546A-27/c
Sequence 27, Application US/08488546A
Patent No. 6054313
GENERAL INFORMATION:
APPLICANT: Joseph Bryan, Lydia Aguilar Bryan, Daniel Nelson, Pamela
APPLICANT: Thomas, Gilbert Cote, and Robert Gagel
TITLE OF INVENTION: Sequence Encoding Mammalian Sulfonyleurea Receptor
Patent No. 6054313
NUMBER OF SEQUENCES: 49
CORRESPONDENCE ADDRESS:
ADDRESS: Woodcock Washburn Kurtz Mackiewicz &
ADDRESSEE: No. 6054313r's
STREET: One Liberty Place 46th. Floor
CITY: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19103

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/488,546A
FILING DATE: 07-JUNE-1995
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/404,531
FILING DATE: 15-MARCH-1995
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Beardell, Lori Y.
REGISTRATION NUMBER: 34,293
REFERENCE/DOCKET NUMBER: BYLR-0026
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-568-3100
TELEFAX: 215-568-3439
INFORMATION FOR SEQ ID NO: 27:
SEQUENCE CHARACTERISTICS:
LENGTH: 4635 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHEICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
FEATURE:
NAME/KEY: CDS
LOCATION: 37..4533

```

RESULT 8
US-08-809-254A-4/C
Sequence 4, Application US/08809254A
Patent No. 6660852
GENERAL INFORMATION:
APPLICANT: KESHI ET AL
TITLE OF INVENTION: PROBE FOR DIAGNOSING INFECTIOUS DISEASES
FILE REFERENCE: 19036/33/67
CURRENT APPLICATION NUMBER: US/08/809,254A
CURRENT FILING DATE: 1997-05-16
PRIORITY FILING DATE: 1995-10-02
PRIORITY APPLICATION NUMBER: JP 236348
PRIORITY FILING DATE: 1994-09-30
NUMBER OF SEQ ID NOS: 9
SOFTWARE: PatentIn version 3.1
SEQ ID NO 4
LENGTH: 5829
TYPE: DNA
ORGANISM: Artificial sequence
FEATURES:
OTHER INFORMATION: Synthetic probe
US-08-809-254A-4

Query Match 9.5%; Score 30.8; DB 4; Length 5829;
Best Local Similarity 57.1%; Pred. No. 1.4;
Matches 56; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

QY 122 AATGGCGATGTGATGCGCAGAGATCCTTGCATTTGACCGCATCAAGTTGCA 181
DB 671 ATATGGCAGAAATTTCAATCAGTTATCTTGCACCTTCTCGCATATTTAAAGGA 612
QY 182 GAATGCGATAGATCTGAAGACTCAACCTTGTGCTC 219
DB 611 AAGATGCTAGAGATATGGAATATCGTCCTTTGTATC 574

RESULT 9
US-08-232-463-14
Sequence 14, Application US/08232463
Patent No. 5670367
GENERAL INFORMATION:
APPLICANT: DORNER, F.
APPLICANT: SCHEIFLINGER, F.
APPLICANT: FALKNER, F. G.
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSER: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,463
FILING DATE:
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US/07/935,313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 IMMU

TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-9300
TELEFAX: (703) 683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: pTZpT-F18
US-08-232-463-14

Query Match 9.4%; Score 30.6; DB 1; Length 7218;
Best Local Similarity 1.0%; Pred. No. 1.9;
Matches 3; Conservative 179; Mismatches 133; Indels 0; Gaps 0;

QY 10 GATTTCCTCTGATCTCTTCTGTTGCGCAGTATGTCAGTCATTTCTTGGCGTGG 69
DB 1064 GATYYY 1123
QY 70 GCGTTCGCTTCTTAATCCTTAAGCCCTAGCATGATGACATTTGGCAGAAATGSCG 129
DB 1124 YY 1183
QY 130 ATGTGATGCGCAGAGATCCTTGCATTTGACCGCATCAAGCTTGGCAGATGCG 189
DB 1184 YY 1243
QY 190 ATAGATCTGAGACCTTCAACCTTGTGCTCATGTCGCTGAGTACAGTGGTGGCAT 249
DB 1244 YY 1303
QY 250 CTCTTGGCGGTGAGCAGCACCCTTGTGAGCGTGGCGCTGGCACCATGTTTGTGA 309
DB 1304 YY 1363
QY 310 CTGTGATCTTTC 324
DB 1364 YYYYYYYYYYYYYY 1378

RESULT 10
US-08-404-531B-27/C
Sequence 27, Application US/08404531B
Patent No. 5863724
GENERAL INFORMATION:
APPLICANT: Joseph Bryan, Lydia Aguilar Bryan, Daniel Nelson, Pamela
APPLICANT: Thomas, Gilbert Cole, and Robert Gagel
TITLE OF INVENTION: Sequence Encoding Mammalian Sulfonylurea Receptor
Patent No. 5863724
NUMBER OF SEQUENCES: 49
CORRESPONDENCE ADDRESS:
ADDRESSER: Woodcock Washburn Kurtz Mackiewicz &
ADDRESSER: No. 5863724is
STREET: One Liberty Place 46th. Floor
CITY: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19103
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/404,531B
FILING DATE: 15-MAR-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Beardsell, Lori Y.
REGISTRATION NUMBER: 34,293

PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 13729
LENGTH: 594
TYPE: DNA
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-13729

Query Match
Best Local Similarity 56.8%; Pred. No. 0.023;
Matches 63; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

QY 204 TCTAACCCTTGTCTGATCCGCTGCGATTCAGTGTGGGCGCATCTTTGGGGGTG 263
DB 407 TATAGGCTTGGCCCTGCGTGGAGTCCGCGGGCGGCTTGAATCTTGCTGATCC 348
QY 264 AGCACCCTTGTGAGCGGTGGCGGCACATCGTTTGTGTGACTGCT 314
DB 347 GCGCCAGGGGCTTCTGATCGGGGTCTCTCGCGCTTGTGCGGACAGCT 297

RESULT 5
US-09-252-991A-13832/C
Sequence 13832, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
PRIOR FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 13832
LENGTH: 678
TYPE: DNA
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-13832

Query Match
Best Local Similarity 56.8%; Pred. No. 0.025;
Matches 63; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

QY 204 TCTAACCCTTGTCTGATCCGCTGCGATTCAGTGTGGGCGCATCTTTGGGGGTG 263
DB 472 TATAGGCTTGGCCCTGCGTGGAGTCCGCGGGCGGCTTGAATCTTGCTGATCC 413
QY 264 AGCACCCTTGTGAGCGGTGGCGGCACATCGTTTGTGTGACTGCT 314
DB 412 GCGCCAGGGGCTTCTGATCGGGGTCTCTCGCGGCTTGTGCGGACAGT 362

RESULT 6
US-09-328-352-1339
Sequence 1339, Application US/09328352
Patent No. 6562958
GENERAL INFORMATION:
APPLICANT: Gary L. Breton et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
FILE REFERENCE: BAUMANNIT FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: GTC99-03PA
CURRENT APPLICATION NUMBER: US/09/328,352
CURRENT FILING DATE: 1999-06-04
NUMBER OF SEQ ID NOS: 8252
SEQ ID NO 1339
LENGTH: 1404
TYPE: DNA
ORGANISM: Acinetobacter baumannii

US-09-328-352-1339

Query Match
Best Local Similarity 55.5%; Pred. No. 0.32;
Matches 61; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

QY 136 ATGCCAGGAGATCCCTGCGATTTGACCGCATCAACGTTTGGCAGCATGATAGAT 195
DB 541 ATTCCAATGGAGATATGGCATATTAATGACATTAAGTCAATGCCAAATGAAGG 600
QY 196 CTGAAGACTTAACCTTTGTCTCATTTCCGTTGCGATTACAGTGTGC 245
DB 601 CTTACGTTAACTTTGATCTTATGATTTGATTTGTCTCTTACGTCTGCG 650

RESULT 7
US-08-961-527-56
Sequence 56, Application US/08961527
Patent No. 6420135
GENERAL INFORMATION:
APPLICANT: Charles Kunsch
TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
CORRESPONDENCE ADDRESS:
NUMBER OF SEQUENCES: 391
ADDRESS: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4MB storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/961,527
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Brookes, A. Anders
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PB340P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8512
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 56:
SEQUENCE CHARACTERISTICS:
LENGTH: 2387 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-961-527-56

Query Match
Best Local Similarity 52.7%; Pred. No. 0.51;
Matches 68; Conservative 0; Mismatches 61; Indels 0; Gaps 0;

QY 112 TTTGTGGCAAAAGGCGATGTGATGCCAGCAAGAACTTGGCAATTTGACCGCATCA 171
DB 962 TTTGTGAAAAAAGACATTAATTTGACAGAAAGCAAAATTAAGTTAAATGCTTTA 1021
QY 172 AGCTTCCAGCAATGATGATCTGAAGCTTAACCTTATGCTCATTCGCGGTGCG 231
DB 1022 GCAATTCAGCGTATTTCCAAATGAGAAATTAACAGTTAGTCTTTCCTTAAT 1081
QY 232 ATTAAGTG 240
DB 1082 ATGATGTG 1090

Qy	241	GTGGCGCATCTTTCTTGGCGGTCGACCACTTTGTGAAGCGTGGCGTGACCATCGTT	300
Db	241	GTGGCGCATCTTTCTTGGCGGTCGACCACTTTGTGAAGCGTGGCGTGACCATCGTT	300
Qy	301	TTTGTGGACTGGGGAATCTTTTC	324
Db	301	TTTGTGGACTGGGGAATCTTTTC	324

RESULT 2
US-09-471-803A-1
; Sequence 1, Application US/09471803A
; Date of Invention 09/15/09

```

1 GENERAL INFORMATION:
2 APPLICANT: KENNERKNECHT, NICOLE
3 APPLICANT: SAHM, HERMANN
4 APPLICANT: EGSELING, LOTMAR
5 APPLICANT: PFEFFERLE, WALTER
6 TITLE OF INVENTION: NUCLEOTIDE SEQUENCE CODING FOR THE EXPORT OF
7 TITLE OF INVENTION: BRANCHED CHAIN AMINO ACIDS, PROCESS FOR THE
8 TITLE OF INVENTION: ISOLATION THEREOF AND USE THEREOF
9 FILE REFERENCE: 21123/265496/MA5
10 CURRENT APPLICATION NUMBER: US/09/471, 803A
11 CURRENT FILING DATE: 1999-12-23
12 PRIOR APPLICATION NUMBER: DE 199 51 708.8
13 PRIOR FILING DATE: 1999-10-27
14 NUMBER OF SEQ ID NOS: 12
15 SOFTWARE: Patentin Ver. 2.1
16 SEQ ID NO 1
17 LENGTH: 1271
18 TYPE: DNA
19 ORGANISM: Corynebacterium glutamicum
20 FEATURE:
21 NAME/KEY: gene
22 LOCATION: (101)..(853)
23 OTHER INFORMATION: brnF
24 NAME/KEY: gene
25 LOCATION: (853)..(1176)
26 OTHER INFORMATION: brnE
27 OTHER INFORMATION: ATCC14752

```

Query Match	100.0%	Score 324	DB 4	Length 1271
Best Local Similarity	100.0%	Pred	0.5e-107	
Matches 324	Conservative 0	Mismatches 0	Indels 0	Gaps 0

Qy	1	ANGACAACGATTTCTCCGTAATTCCTTTGGTCGAGATGTGCAGTATCACTACTTTT	60
Dp	853	ATGACAACGATTTCTCCGTAATTCCTTTGGTCGAGATGTGCAGTATCACTACTTTT	912
Qy	61	GGCGTCGGGGCGGTCCTCTTAAATCCTTAAGCCCCTAACGGAATACAAATTTGTGGGC	120
Dp	913	GGCGTCGGGGCGGTCCTCTTAAATCCTTAAGCCCCTAACGGAATACAAATTTGTGGGC	972
Qy	121	AAATATGCGCATGTGATGCGACAGAGAAATCTTGCATTTTGAACCGCATCAACTGTTGCG	180
Dp	973	AAATATGCGCATGTGATGCGACAGAGAAATCTTGCATTTTGAACCGCATCAACTGTTGCG	1032
Qy	181	AGCAATGCGATGATCTGAAGACTCTAACCTTTGGTCTCATTTGCCGTTGCCATTTACAGTG	240
Dp	1033	AGCAATGCGATGATCTGAAGACTCTAACCTTTGGTCTCATTTGCCGTTGCCATTTACAGTG	1092
Qy	241	GTGGCGCATCTCTTGGCGGCTGACAGCAACTTTGAGCGTTGCGCGCTGACCACTCGTT	300
Dp	1093	GTGGCGCATCTCTTGGCGGCTGACAGCAACTTTGAGCGTTGCGCGCTGACCACTCGTT	1152
Qy	301	TTTGTGGACTGGTGAATCTTTTC	324
Dp	1153	TTTGTGGACTGGTGAATCTTTTC	1176

RESULT 3
US-09-471-803A-6

Sequence 6, Application US/094771803A
Patent No. 6613545
GENERAL INFORMATION:
APPLICANT: KENNERKNECHT, NICOLE
APPLICANT: SAMM, HERMANN
APPLICANT: EGSELING, LOTHAR
APPLICANT: PFEFFEL, WALTER
TITLE OF INVENTION: NICOTINIDE SEQUENCE CODING FOR THE EXPORT OF
TITLE OF INVENTION: BRANCHED CHAIN AMINO ACIDS, PROCESS FOR THE
TITLE OF INVENTION: ISOLATION THEREOF AND USE THEREOF
FILE REFERENCE: 21123/265496/MAS
CURRENT APPLICATION NUMBER: US/09/471,803A
CURRENT FILING DATE: 1999-12-23
PRIOR APPLICATION NUMBER: DE 199 51 708.8
PRIOR FILING DATE: 1999-10-27
NUMBER OF SEQ ID NOS: 12
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO. 6

OTHER INFORMATION: ATCC13032
US-09-471-803A-6

Query Match	100.0%	Score 324	DB 4	Length 1271
Best Local Similarity	100.0%	Pred. No. 5e-107		
Best Match 324	Conservative	0	Mismatches	0
			Indels	0
			Gaps	0

QY	1	ATGCAACTGATATTTCTCTCTGATATCTCTTTTGTGCGAGATGAGCAGTCACTTACTTT	60
Db	853	ATGCAACTGATATTTCTCTCTGATATCTCTTTTGTGCGAGATGAGCAGTCACTTACTTT	912
QY	61	GGCGTCCGGGGCGGTCCTCGTTCTTAATCCTTAAGCCCTCACTGAATCACAATTTGTGGGC	120
Db	913	GGCGTCCGGGGCGGTCCTCGTTCTTAATCCTTAAGCCCTCACTGAATCACAATTTGTGGGC	972
QY	121	AAATGGCGAGTGTGATATCCAGCAGAGATCCTTGCATTTTGTAGCCGCATCAAGCTTTGGC	180
Db	973	AAATGGCGAGTGTGATATCCAGCAGAGATCCTTGCATTTTGAOCGATCAACGTTTGGC	1032
QY	181	AGCAATGCGATAGATCTGGAAGACTCTTAACCTTTGGTGTCAATGGCGCTTGGCATTAACGTG	240
Db	1033	AGCAATGCGATAGATCTGGAAGACTCTTAACCTTTGGTGTCAATGGCGCTTGGCATTAACGTG	1092
QY	241	GTGGCGCATCTTCTTGTGGCGGTGCAGCGACCTGTGTGAGCGTTGGCGCTG3GACCAATCGTT	300
Db	1093	GTGGCGCATCTTCTTGTGGCGGTGCAGCGACCTGTGTGAGCGTTGGCGCTG3GACCAATCGTT	1152
QY	301	TTTGTGTGAGCTGTGGAATCTTTTC	324
Db	1153	TTTGTGTGAGCTGTGGAATCTTTTC	1176

RESULT 4
US-09-252-991A-13729/C
; Sequence 13729, Application US/09252991A

1 APPLICANT: Marc J. Rudenfield et al.
2 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
3 TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
4 FILE REFERENCE: 107196.136
5 CURRENT APPLICATION NUMBER: US/09/252,991A
6 CURRENT FILING DATE: 1999-02-18
7 PRIOR APPLICATION NUMBER: US 60/074,788
8 PRIOR FILING DATE: 1998-02-18

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 19, 2004, 10:57:33 ; Search time 27.2807 Seconds
(without alignments)
6590.898 Million cell updates/sec

Title: US-10-608-504-6_COPY_853_1176

Perfect score: 324
Sequence: 1 atgacacactgattctctctg.....ttgactgctgactctcttc 324

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: Issued Patents NA:
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfillseq1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	324	100.0	324	4 US-09-471-803A-4	Sequence 4, Appl
2	324	100.0	1271	4 US-09-471-803A-1	Sequence 1, Appl
3	324	100.0	1271	4 US-09-471-803A-6	Sequence 1, Appl
4	34.2	10.6	594	4 US-09-252-991A-13729	Sequence 13729, A
5	34.2	10.6	678	4 US-09-252-991A-13832	Sequence 13832, A
6	31.6	9.8	1404	4 US-09-328-352-1339	Sequence 1339, Ap
7	31.4	9.7	2387	4 US-08-961-527-56	Sequence 56, Appl
8	30.8	9.5	5829	4 US-08-809-254A-4	Sequence 4, Appl
9	30.6	9.4	7218	1 US-08-232-463-14	Sequence 14, Appl
10	29.8	9.2	4635	2 US-08-404-531B-27	Sequence 27, Appl
11	29.8	9.2	4635	3 US-08-476-900A-27	Sequence 27, Appl
12	29.8	9.2	4635	3 US-08-488-546A-27	Sequence 27, Appl
13	29.8	9.2	5110	2 US-08-404-531B-5	Sequence 4, Appl
14	29.8	9.2	5110	2 US-08-404-531B-5	Sequence 5, Appl
15	29.8	9.2	5110	3 US-08-476-900A-4	Sequence 4, Appl
16	29.8	9.2	5110	3 US-08-476-900A-5	Sequence 5, Appl
17	29.8	9.2	5110	3 US-08-488-546A-4	Sequence 4, Appl
18	29.8	9.2	5110	3 US-08-488-546A-5	Sequence 5, Appl
19	29.6	9.1	1308	2 US-08-404-531B-1	Sequence 1, Appl
20	29.6	9.1	1308	2 US-08-404-531B-2	Sequence 2, Appl
21	29.6	9.1	1308	2 US-08-476-900A-1	Sequence 1, Appl
22	29.6	9.1	1308	3 US-08-476-900A-2	Sequence 2, Appl
23	29.6	9.1	1308	3 US-08-488-546A-1	Sequence 1, Appl
24	29.6	9.1	1308	3 US-08-488-546A-2	Sequence 2, Appl
25	29.6	9.1	2454	2 US-08-404-531B-32	Sequence 32, Appl
26	29.6	9.1	2454	3 US-08-476-900A-32	Sequence 32, Appl
27	29.6	9.1	2454	3 US-08-488-546A-32	Sequence 32, Appl

28	29.6	9.1	4931	3 US-08-726-320-2	Sequence 2, Appl
29	29.6	9.1	4931	3 US-09-208-716-2	Sequence 2, Appl
30	29.4	9.1	611	4 US-09-976-594-152	Sequence 152, App
31	29.4	9.1	12588	2 US-08-387-942C-1	Sequence 1, Appl
32	29.2	9.0	867	4 US-09-216-393B-340	Sequence 340, App
33	29.2	9.0	867	4 US-09-216-393B-342	Sequence 342, App
34	29.2	9.0	1397	4 US-09-216-393B-343	Sequence 343, App
35	29.2	9.0	1397	4 US-09-216-393B-345	Sequence 345, App
36	29.2	9.0	9828	4 US-08-961-527-41	Sequence 41, Appl
37	28.8	8.9	455	2 US-09-621-976-15639	Sequence 15639, A
38	28.4	8.8	459	2 US-08-387-942C-35	Sequence 35, Appl
39	28.4	8.8	7304	4 US-09-453-702B-174	Sequence 174, App
40	28.2	8.7	387	4 US-08-651-155B-74	Sequence 74, Appl
41	28.2	8.7	387	4 US-08-194-036B-74	Sequence 74, Appl
42	28.2	8.7	5099	4 US-09-610-040-5	Sequence 5, Appl
43	28.2	8.7	1230025	4 US-09-198-452A-1	Sequence 1, Appl
44	28	8.6	531	4 US-08-671-548C-7	Sequence 7, Appl
45	28	8.6	1758	4 US-09-205-258-173	Sequence 173, App

ALIGNMENTS

RESULT 1
US-09-471-803A-4
Sequence 4, Application US/09471803A

Patent No. 6613545
GENERAL INFORMATION:

APPLICANT: KENNERKNECHT, NICOLE
APPLICANT: SAHM, HERMANN

APPLICANT: EGGELING, LOTMAR
APPLICANT: PFEFFERLE, WALTER

TITLE OF INVENTION: NUCLEOTIDE SEQUENCE CODING FOR THE EXPORT OF
TITLE OF INVENTION: BRANCHED CHAIN AMINO ACIDS, PROCESS FOR THE

FILE REFERENCE: 21123/265496/MAS
CURRENT FILING DATE: 1999-12-23

PRIOR APPLICATION NUMBER: DE 199 51 708.8
PRIOR FILING DATE: 1999-10-27

NUMBER OF SEQ ID NOS: 12
SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 4
LENGTH: 324

TYPE: DNA

ORGANISM: Corynebacterium glutamicum
FEATURE:

NAME/KEY: CDS
LOCATION: (1)..(324)

OTHER INFORMATION: brne
OTHER INFORMATION: ATCC14752

US-09-471-803A-4

Query Match
Best Local Similarity 100.0%; Pred. No. 2.4e-107; Length 324;
Matches 324; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	ATGACACTGATTCTCTGATTCTCTGTTGTCGAGTATGTCAGTACTTATCTTT	60
DB	1	ATGACACTGATTCTCTGATTCTCTGTTGTCGAGTATGTCAGTACTTATCTTT	60
QY	61	GCGCTCCGGGGGTTCCGTTCTTAATCTTAAGCCCTTACGTAATCATTTGTGGGC	120
DB	61	GCGCTCCGGGGGTTCCGTTCTTAATCTTAAGCCCTTACGTAATCATTTGTGGGC	120
QY	121	AAATGGGAGTGGATCCAGAGATCTTCCATTGACCGCATCACTTTGGC	180
DB	121	AAATGGGAGTGGATCCAGAGATCTTCCATTGACCGCATCACTTTGGC	180
QY	181	AGCAATGGATGATCTTAAGACTCTAACCTTTGGTCTCATTCGCTTGCATTACGTG	240
DB	181	AGCAATGGATGATCTTAAGACTCTAACCTTTGGTCTCATTCGCTTGCATTACGTG	240

CC The present invention relates to oligonucleotide clones originating in
CC barley (Hordeum vulgare) which contain single nucleotide polymorphisms
CC (SNP). The oligonucleotides may be used for analysis of SNPs among barley
CC varieties, identification of particular varieties and genotype-phenotype
CC analysis, isolation of specific genes and creation of new varieties by
CC transformation of barley varieties with them and production of new barley
CC varieties with desired properties. The present sequence represents an
CC oligonucleotide clone DNA sequence featured in the specification. The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published-pct-sequences. (Updated on 27-OCT-2003 to
CC standardise OS field)

XX Sequence 470 BP; 122 A; 131 C; 102 G; 106 T; 0 U; 9 Other;

Query Match 9.9%; Score 32; DB 8; Length 470;

Best Local Similarity 51.9%; Pred. No. 2.2;
Matches 68; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

QY	181	ACCAATGGCATAGATCTGAAGACTCTAACCTTTGCTCATGGCCGTTGGATTACAGTG	240
DB	384	ATCAAGNACAGGAGNAGCATTAAGGAACTTTGAGNCCATCCACGTTGAGAGATGCCAG	325
QY	241	GTGGGGCATCTTTGGGGGTGACGCACTTTGTTAGCGTTGGCGCTGGCACCATCGTT	300
DB	324	GGTACACAGTTTGGCAACCCGCTGGGCAAGCTGTTCACATTGGCAAGGGCAACCAAGCT	265
QY	301	TTGTTGACT	311
DB	264	TGGGTGAGCT	254

Search completed: April 19, 2004, 11:42:06
Job time : 151.886 secs

CC The present invention relates to oligonucleotide clones originating in
CC barley (*Hordeum vulgare*) which contain single nucleotide polymorphisms
CC (SNP). The oligonucleotides may be used for analysis of SNPs among barley
CC varieties, identification of particular varieties and genotype-phenotype
CC analysis, isolation of specific genes and creation of new varieties by
CC transformation of barley varieties with them and production of new barley
CC varieties with desired properties. The present sequence represents an
CC oligonucleotide clone DNA sequence featured in the specification. The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at <http://wipo.int/pub/published-pct-sequences>. (Updated on 27-OCT-2003 to
CC standardise OS field)
XX
SQ Sequence 376 BP; 97 A; 105 C; 82 G; 92 T; 0 U; 0 Other;
Query Match 9.9%; Score 32.2; DB 8; Length 376;
Best Local Similarity 53.6%; Pred. No. 1.7; Mismatches 0; Gaps 0;
Matches 67; Conservative 0; Indels 58; Indels 0; Gaps 0;
QY 187 GCGATGATCTGAAGACTCTAACCTTTGCTCATTTGCCGTTGCGATTACAGTGTGGCG 246
DB 375 GAGAGAGAGAGAGCATTAAGGAACTTTGAGACCATCCAGTTCAGGATGCCAGGGTCAC 316
QY 247 CATCTTTTGGCGGTCACGACCACTTTGAGCGTTGGCGCTGGCAACCATGTTTGT 306
DB 315 CAGTTTGGCAACCCGCTGTGGGCAACGTTTCACCATTTGGCAAGGACCAAGCCTTGGGTG 256
QY 307 GGAAT 311
DB 255 AGCCT 251
RESULT 14
ACL20868/c
ID ACL20868 standard; DNA; 392 BP.
XX
AC ACL20868;
XX
DT 27-OCT-2003 (revised)
DT 17-OCT-2003 (first entry)
XX
DE DNA clone originating in barley containing SNP encoding sequence #10859.
XX
KM Barley; single nucleotide polymorphism; SNP; genotype-phenotype analysis;
KM gene; ss.
XX
OS *Hordeum vulgare*; var. (cul. Akashinriki).
XX
PN WO2003057877-A1.
XX
PD 17-JUL-2003.
XX
PF 16-DEC-2002; 2002WO-IB005403.
XX
PR 20-DEC-2001; 2001JP-00387059.
PR 20-DEC-2001; 2001JP-00387131.
PR 20-DEC-2001; 2001JP-00403299.
PR 20-DEC-2001; 2001JP-00403300.
PR 27-SEP-2002; 2002JP-00327515.
XX
PA (UNNT-) UNIV JAPAN OKAYAMA.
XX
PI Sato K, Takeda K, Kohara Y;
XX
DR WPI; 2003-587127/55.
XX
PT Single nucleotide polymorphism sites in barley varieties and DNA
PT sequences containing them for analysis and identification of barley
PT varieties and production of barley transformants with desired
PT characteristics.
XX
PS Disclosure; SEQ ID XX; 284bp; Japanese.
XX

CC The present invention relates to oligonucleotide clones originating in
CC barley (*Hordeum vulgare*) which contain single nucleotide polymorphisms
CC (SNP). The oligonucleotides may be used for analysis of SNPs among barley
CC varieties, identification of particular varieties and genotype-phenotype
CC analysis, isolation of specific genes and creation of new varieties by
CC transformation of barley varieties with them and production of new barley
CC varieties with desired properties. The present sequence represents an
CC oligonucleotide clone DNA sequence featured in the specification. The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at <http://wipo.int/pub/published-pct-sequences>. (Updated on 27-OCT-2003 to
CC standardise OS field)
XX
SQ Sequence 392 BP; 100 A; 105 C; 85 G; 102 T; 0 U; 0 Other;
Query Match 9.9%; Score 32.2; DB 8; Length 392;
Best Local Similarity 53.6%; Pred. No. 1.7; Mismatches 0; Gaps 0;
Matches 67; Conservative 0; Indels 58; Indels 0; Gaps 0;
QY 187 GCGATGATCTGAAGACTCTAACCTTTGCTCATTTGCCGTTGCGATTACAGTGTGGCG 246
DB 391 GAGAGAGAGAGAGCATTAAGGAACTTTGAGACCATCCAGTTCAGGATGCCAGGGTCAC 332
QY 247 CATCTTTTGGCGGTCACGACCACTTTGAGCGTTGGCGCTGGCAACCATGTTTGT 306
DB 331 CAGTTTGGCAACCCGCTGTGGGCAACGTTTCACCATTTGGCAAGGACCAAGCCTTGGGTG 272
QY 307 GGAAT 311
DB 271 AGCCT 267
RESULT 15
ACL20918/c
ID ACL20918 standard; DNA; 470 BP.
XX
AC ACL20918;
XX
DT 27-OCT-2003 (revised)
DT 17-OCT-2003 (first entry)
XX
DE DNA clone originating in barley containing SNP encoding sequence #10909.
XX
KM Barley; single nucleotide polymorphism; SNP; genotype-phenotype analysis;
KM gene; ss.
XX
OS *Hordeum vulgare*; var. (cul. Haruna Nijo).
XX
PN WO2003057877-A1.
XX
PD 17-JUL-2003.
XX
PF 16-DEC-2002; 2002WO-IB005403.
XX
PR 20-DEC-2001; 2001JP-00387059.
PR 20-DEC-2001; 2001JP-00387131.
PR 20-DEC-2001; 2001JP-00403299.
PR 20-DEC-2001; 2001JP-00403300.
PR 27-SEP-2002; 2002JP-00327515.
XX
PA (UNNT-) UNIV JAPAN OKAYAMA.
XX
PI Sato K, Takeda K, Kohara Y;
XX
DR WPI; 2003-587127/55.
XX
PT Single nucleotide polymorphism sites in barley varieties and DNA
PT sequences containing them for analysis and identification of barley
PT varieties and production of barley transformants with desired
PT characteristics.
XX
PS Disclosure; SEQ ID XX; 284bp; Japanese.
XX

Query Match 10.0%; Score 32.4; DB 6; Length 6672;
 Best Local Similarity 58.6%; Pred. No. 5.4;
 Matches 75; Conservative 0; Mismatches 51; Indels 2; Gaps 1;

QY 93 GCCCCTACGTAATCAATTTGTGGGCAAAATGCGCATGTGATGCCAGCAGAAATCCT 152
 DB 5382 GCACACACGTCATGTGCACATGTGTGAAAGAGGCTGGTGAAGACATGCAATCCT 5323
 QY 153 TGGCATTTTGACCGCATCAACGTTTGGCAGCAATGGAGTATGATCTGAAAGCTTAACCTT 212
 DB 5322 TGTCACTGTC--CAACAGAAATGTGCAGAAATGGAGTGGCTGGTGTGTGTACACCTG 5265
 QY 213 TGGTCTCA 220
 DB 5264 TAATCCCA 5257

RESULT 12
 ACF39600/C
 ID ACF39600 standard; cDNA; 6672 BP.
 AC ACF39600;
 XX
 XX 29-SEP-2003 (first entry)
 XX
 DE Human MHC class II transactivator encoding cDNA SEQ ID NO:3.
 XX
 XX Human; major histocompatibility complex class II transactivator;
 KM MHC class II transactivator; antisense modulation; immunosuppressive;
 KM antimicrobial; antidiabetic; antirheumatic; antitachytic; cycostatic;
 KM neurotropic; neuroprotective; immunostimulant; autoimmune disorder;
 KM MHC class II transactivator; inhibitor; infection; transplant rejection;
 KM diabetes; rheumatoid arthritis; cancer; Alzheimer's disease;
 KM multiple sclerosis; severe combined immunodeficiency disease; gene; ss.
 XX
 OS Homo sapiens.
 XX
 XX Key Location/Qualifiers
 FT 139..3531
 FT CDS /*tag= a
 FT /product= "MHC class II transactivator"
 FT
 FN MO2003050247-A2.
 PD 19-JUN-2003.
 PF 04-DEC-2002; 2002MO-US038616.
 XX
 XX 05-DEC-2001; 2001US-00006366.
 XX
 XX (ISIS-) ISIS PHARM INC.
 XX
 XX Bennett FC, Dobie KM;
 PI
 DR WPI; 2003-577294/54.
 DR P-PSDB; ABR82108.
 XX
 PT New antisense oligonucleotides for modulating MHC class II transactivator
 PT gene expression, particularly useful for treating autoimmune disorders
 PT such as transplant rejection, Alzheimer's disease, or multiple sclerosis,
 PT or infection.
 PT
 XX
 XX Example 13; Page 93-99; 129pp; English.
 XX
 XX The present invention describes a compound (I) that is 8-50 nucleobases
 CC in length; (a) targets a nucleic acid molecule encoding major
 CC histocompatibility complex (MHC) class II transactivator, and
 CC specifically hybridises with the nucleic acid encoding the MHC class II
 CC transactivator, and inhibits the expression of MHC class II
 CC transactivator; or (b) specifically hybridises with at least an 8-
 CC nucleobase portion of an active site on a nucleic acid molecule encoding
 CC MHC class II transactivator. (I) has immunosuppressive, antimicrobial,
 CC antidiabetic, antirheumatic, antitachytic, cycostatic, neurotropic,

CC neuroprotective and immunostimulant activities, and can be used as an MHC
 CC class II transactivator inhibitor. The MHC class II transactivator
 CC antisense oligonucleotides can be used for treating an animal having a
 CC disease or condition associated with MHC class II transactivator, e.g.,
 CC autoimmune disorder or infection. The antisense oligonucleotides can be
 CC used for inhibiting the expression of MHC class II transactivator in
 CC cells or tissues. In particular, these diseases include transplant
 CC rejection, diabetes, rheumatoid arthritis, cancer, Alzheimer's disease,
 CC multiple sclerosis, or severe combined immunodeficiency disease. The
 CC antisense compounds are useful for diagnostics, prophylaxis, or as
 CC research reagents or kits. The present sequence encodes a human MHC class
 CC II transactivator, which is used in an example from the present invention
 XX
 XX Sequence 6672 BP; 1392 A; 1954 C; 1868 G; 1458 T; 0 U; 0 Other;
 XX

Query Match 10.0%; Score 32.4; DB 8; Length 6672;
 Best Local Similarity 58.6%; Pred. No. 5.4;
 Matches 75; Conservative 0; Mismatches 51; Indels 2; Gaps 1;

QY 93 GCCCCTACGTAATCAATTTGTGGGCAAAATGCGCATGTGATGCCAGCAGAAATCCT 152
 DB 5382 GCACACACGTCATGTGCACATGTGTGAAAGAGGCTGGTGAAGACATGCAATCCT 5323
 QY 153 TGGCATTTTGACCGCATCAACGTTTGGCAGCAATGGAGTATGATCTGAAAGCTTAACCTT 212
 DB 5322 TGTCACTGTC--CAACAGAAATGTGCAGAAATGGAGTGGCTGGTGTGTGTACACCTG 5265
 QY 213 TGGTCTCA 220
 DB 5264 TAATCCCA 5257

RESULT 13
 ACL20925/C
 ID ACL20925 standard; DNA; 376 BP.
 AC ACL20925;
 XX
 XX 27-OCT-2003 (revised)
 DT 17-OCT-2003 (first entry)
 XX
 DE DNA clone originating in barley containing SNP encoding sequence #10916.
 XX
 KM Barley; single nucleotide polymorphism; SNP; genotype-phenotype analysis;
 KM gene; ss.
 XX
 OS Hordeum vulgare; var. (cul. Akashinrkt).
 XX
 XX WO2003057877-A1.
 XX
 XX 17-JUL-2003.
 PD
 PF 16-DEC-2002; 2002MO-IB005403.
 XX
 XX 20-DEC-2001; 2001JP-00387059.
 PR 20-DEC-2001; 2001JP-00387131.
 PR 20-DEC-2001; 2001JP-00403299.
 PR 20-DEC-2001; 2001JP-00403300.
 PR 27-SEP-2002; 2002JP-00327515.
 XX
 XX (UNM-) UNIV JAPAN OKAYAMA.
 XX
 XX Sato K, Takeda K, Kohara Y;
 PI
 DR WPI; 2003-587127/55.
 XX
 PT Single nucleotide polymorphism sites in barley varieties and DNA
 PT sequences containing them for analysis and identification of barley
 PT varieties and production of barley transformants with desired
 PT characteristics.
 PT
 XX Disclosure; SEQ ID XX; 284bp; Japanese.
 XX

CC capable of detecting 1000 or more genes from *Drosophila*. The invention is
CC useful in developmental biology and in elucidating cell signaling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (AB116176-AB130511), expressed DNA
CC sequences (AB101840-AB116175) and the encoded proteins (AB57237-
CC AB572072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIP0 at ftp.wipo.int/pub/published_pct_sequences

XX Sequence 3144 BP; 1056 A; 676 C; 680 G; 732 T; 0 U; 0 Other;

Query Match 10.1%; Score 32.8; DB 4; Length 3144;
Best Local Similarity 52.1%; Pred. No. 2.8; Mismatches 0; Gaps 0;
Matches 73; Conservative 0; Indels 67; Indels 0; Gaps 0;

169 TCACGTTTGGCAGCATGCGATGATCTGAGACTCTACCTTGGTCATTCGCTT 228
1512 TCCTAATTTCTATATATATTGTTAAATGATGTTATTAATGTTACCTGATATCTGTT 1453

229 GCGATTACAGTGTGGCGCATCTTCTTGGCGTCGACGACCTGTTGAGCGTGGCGCT 288

1452 TCGTTTGGGGTGTGTGTTTCTTTCGTCGCGGACGACATGCTTGTGCTTGGCTGCA 1393

289 GGCACCATCGTTTGTGTTGG 308

1392 CACAGTCTTATGTTGTTGG 1373

RESULT 9
AAS59515/c
ID AAS59515 standard; DNA; 66788 BP.

XX AAS59515;

XX 13-FEB-2002 (first entry)

XX Propionibacterium acnes immunogenic protein encoding DNA #10.

XX SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
XX uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
XX inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
XX dermatological; osteopathic; neuroprotectant; ds.

XX Propionibacterium acnes.

XX WO200181581-A2.

XX 01-NOV-2001.

XX 20-APR-2001; 2001WO-US012865.

XX 21-APR-2000; 2000US-0199047P.

XX 02-JUN-2000; 2000US-0208841P.

XX 07-JUL-2000; 2000US-0216747P.

XX (CORI-) CORIXA CORP.

XX Skeiky YAM, Persing DH, Mitcham JL, Wang SS, Bhattacharya A;

XX L'maisonneuve J, Zhang Y, Jen S, Carter D;

XX WPI; 2001-616774/71.

XX Propionibacterium acnes polypeptides and nucleic acids useful for
XX vaccinating against and diagnosing infections, especially useful for
XX treating acne vulgaris.

XX Claim 1; SEQ ID NO 10; 1069bp; English.

XX Sequences AAS59506-AAS59804 represent DNA molecules encoding
XX propionibacterium acnes immunogenic polypeptides. The proteins and their
XX associated DNA sequences are used in the treatment, prevention and
XX diagnosis of medical conditions caused by *P. acnes*. The disorders include

CC SAPHO syndrome (synovitis, acne, pustulosis, hypertosis and
CC osteomyelitis), uveitis and endophthalmitis. *P. acnes* is also involved in
CC infections of bone, joints and the central nervous system, however it is
CC particularly involved in the inflammatory lesions associated with acne
CC vulgaris. A method for detecting the presence or absence of *P. acnes* in a
CC patient comprises contacting a sample with a binding agent that binds to
CC the proteins of the invention and determining the amount of bound protein
CC in the sample. The polypeptides may be used as antigens in the production
CC of antibodies specific for *P. acnes* proteins. These antibodies can be
CC used to downregulate expression and activity of *P. acnes* polypeptides and
CC therefore treat *P. acnes* infections. The antibodies may also be used as
CC diagnostic agents for determining *P. acnes* presence, for example, by
CC enzyme linked immunosorbent assay (ELISA). This sequence encodes the
CC polypeptides shown in AAU1313-42019 and AAU67473. Note: The sequence
CC data for this patent did not form part of the printed specification, but
CC was obtained in electronic format directly from WIP0 at
CC ftp.wipo.int/pub/published_pct_sequences

XX Sequence 66788 BP; 13534 A; 21369 C; 19347 G; 12530 T; 0 U; 8 Other;

Query Match 10.1%; Score 32.6; DB 4; Length 66788;
Best Local Similarity 52.6%; Pred. No. 13; Mismatches 64; Indels 0; Gaps 0;
Matches 71; Conservative 0; Indels 64; Indels 0; Gaps 0;

154 GCCATTTTGAACCGCATCAACGTTTGGCAGCATGATCTGAAGACTTAACTTT 213

4085 GTCAATTCGATCGGGGCGACACCGCCGACCTTCTGACCGGCTCTGCTACTGCT 40026

214 GGTCTCATTCGCGGTTGCGATTACAGTGTGGCGCATCTTCTGGCGGTTCACGACCTTG 273

40025 GGGCGTTTGCACCGTCGATGAGTGAAGGCTCCGATTCGATGGCGGCTCAAGATT 39966

274 TTGACGTTGGCGCT 288

39965 CTCAGCTTCAAGCT 39951

RESULT 10
ACF64444/c

ID ACF64444 standard; DNA; 66788 BP.

XX ACF64444;

XX 17-OCT-2003 (first entry)

XX Propionibacterium acnes DNA contig sequence #10.

XX Acne vulgaris; antiseborrheic; dermatological; antibacterial;
XX immunostimulant; immune response; vaccine; ds.

XX Propionibacterium acnes.

XX WO2003033515-A1.

XX 24-APR-2003.

XX 11-OCT-2002; 2002WO-US032727.

XX 15-OCT-2001; 2001US-00978925.

XX (CORI-) CORIXA CORP.

XX Mitcham JL, Skeiky YAM, Persing DH, Bhattacharya A, Maisonneuve JL;

XX Zhang Y, Wang S, Jen S, Lodes MJ, Benson DR, Jones R, Carter D;

XX Bartha B, Valliere-Douglas J;

XX WPI; 2003-381789/36.

XX New Propionibacterium acnes polypeptides and polynucleotides encoding the
XX polypeptide, useful for diagnosing, preventing or treating acne vulgaris,
XX or for stimulating an immune response specific for a *P. acnes* protein.

XX Claim 1, SEQ ID NO 10; 1481bp; English.

XX
XX
The invention relates to an isolated nucleic acid detection reagent

XX

CC The invention relates to an ISO

QY 301 TTTGTTGACTGTGTAATCTTTTC 324
 1153 TTTGTTGACTGTGTAATCTTTTC 1176

RESULT 5
 ID AAH21112 standard; DNA; 1271 BP.
 XX AAH21112;

DT 05-SEP-2001 (first entry)

XX C. glutamicum DNA encoding brnF and brnE.

XX L-amino acid production; brnF; brnE; branched-chain amino acid;
 KM coryneform bacterium; leucine; isoleucine; valine; medicine;
 KM animal nutrition; ds.

XX Corynebacterium glutamicum.

XX Key Location/Qualifiers

FT CDS 101..856
 FT CDS /*tag= a
 FT CDS /product= "brnF"
 FT CDS 853..1179
 FT CDS /*tag= b
 FT CDS /product= "brnE"

XX BP1096010-A1.

XX 02-MAY-2001.

XX 11-OCT-2000; 2000EP-00122057.

XX 27-OCT-1999; 99DE-01051708.

XX (DEGS) DEGUSSA AG.

XX (KERN) FORSCHUNGSZENTRUM JUELICH GMBH.

XX Kennerknecht N, Eggeling L, Sahm H, Pfeifferle W;

XX WPI; 2001-391595/42.

XX New export genes from coryneform bacteria, useful for increasing

XX fermentative production of branched-chain amino acids.

XX Claim 4 (1); Page 17-18; 23pp; German.

XX This invention describes a novel isolated polynucleotide (1) containing

XX at least one sequence that (1) is 70% identical with a sequence that

XX encodes at least one of 251 amino acid (aa) (3) or 108 aa (5)

XX polypeptides; (ii) encodes a polypeptide at least 70% identical with (3)

XX or (5); (iii) is the complement of (1) or (ii); or (iv) contains at least

XX 15 consecutive bases from (i)-(iii). The invention also describes (a)

XX protein derived from sequences of 753 bp (2) or 324 bp (4); (b)

XX coryneform microorganisms, especially Corynebacterium, transformed with

XX one or more (1), where these are replicative DNA; (c) production of

XX branched-chain L-aa by fermentation of coryneform bacteria in which the

XX brnF and/or brnE genes (or equivalent sequences) are amplified,

XX especially overexpressed; and (d) method for isolating the brnF and/or

XX brnE genes; (1) is used for transformation of coryneform bacteria being

XX used for fermentative production of branched-chain amino acids;

XX specifically leucine, isoleucine and valine, which are useful in medicine

XX and animal nutrition. (1) can also be used as source of primers and

XX probes for isolation of related sequences. Transformation with (1)

XX increases yield of branched-chain amino acids. This sequence encodes the

XX Corynebacterium glutamicum ATCC 13032 brnF and brnE proteins described in

XX the method of the invention

XX Sequence 1271 BP; 265 A; 328 C; 318 G; 360 T; 0 U; 0 Other;

Query Match 100.0%; Score 324; DB 4; Length 1271;
 Best Local Similarity 100.0%; Pred. No. 1.6e-100; Indels 0; Gaps 0;
 Matches 324; Conservative 0; Mismatches 0;

QY 1 ATGACAACTGATTTCTCTGTAATCTTTTC 324
 DB 853 ATGACAACTGATTTCTCTGTAATCTTTTC 1176
 QY 61 GCGCTCCGGGCGGCTTCCTTTAATCTTTTC 120
 DB 913 GCGCTCCGGGCGGCTTCCTTTAATCTTTTC 972
 QY 121 AAATGGCGATGTGATGAGCAGCAGATCTTCCATTGATGACCGCATACGTTTCCG 180
 DB 973 AAATGGCGATGTGATGAGCAGCAGATCTTCCATTGATGACCGCATACGTTTCCG 1032
 QY 181 AGCAATGGATGATCTGAAAGACTCTAATCTTTGCTCATTTGGCGTGGATTAACAGT 240
 DB 1033 AGCAATGGATGATCTGAAAGACTCTAATCTTTGCTCATTTGGCGTGGATTAACAGT 1092
 QY 241 GNGGCGATCTTCTTGGGCGTGCAGCAGCCTTTGTTGACGTTGGCGCTGCACCATCGTT 300
 DB 1093 GNGGCGATCTTCTTGGGCGTGCAGCAGCCTTTGTTGACGTTGGCGCTGCACCATCGTT 1152
 QY 301 TTTGTTGACTGTGTAATCTTTTC 324
 DB 1153 TTTGTTGACTGTGTAATCTTTTC 1176

RESULT 6
 ID AAH64966 standard; DNA; 349980 BP.
 XX AAH64966;

XX 26-SEP-2001 (first entry)

XX C glutamicum coding sequence fragment SEQ ID NO: 1.

XX Coryneform bacterium; amino acid synthesis; vitamin; saccharide;

XX organic acid synthesis; ds.

XX Corynebacterium glutamicum.

XX BP1108790-A2.

XX 20-JUN-2001.

XX 18-DEC-2000; 2000EP-00127688.

XX 16-DEC-1999; 99JP-00377484.

XX 07-APR-2000; 2000JP-00159162.

XX 03-AUG-2000; 2000JP-00280988.

XX (KYOW) KYOWA HAKKO KOGYO KK.

XX Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;

XX Tateishi N, Senoh A, Ikeda W, Ozaki A;

XX WPI; 2001-376931/40.

XX Novel polynucleotides derived from Coryneform bacteria, for identifying

XX mutation point of a gene, measuring expression of a gene, analyzing

XX expression profile or pattern of a gene and identifying homologous gene.

XX Claim 7; SEQ ID NO 1; 246bp + Sequence Listing; English.

XX The present invention provides a number of nucleotide and protein

XX sequences from the Coryneform bacterium Corynebacterium glutamicum. These

XX are useful for identifying the mutation point of a gene derived from a

XX mutant of coryneform bacterium, measuring expression amount and analysing

XX the expression profile or expression pattern of a gene derived from

XX Coryneform bacterium, and identifying a homologue of a gene derived from

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 19, 2004, 10:56:53 ; Search time 150.886 Seconds

(without alignments)
9122.247 Million cell updates/sec

Title: US-10-608-504-6_COPY_853_1176

Sequence: 1 atgacaactgattctctctg.....ctggactggtgactcttcttc 324

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 3373863 seqs, 2124039041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_29Jan04:*
1: geneseqn19808:*
2: geneseqn19908:*
3: geneseqn20008:*
4: geneseqn20018:*
5: geneseqn20018s:*
6: geneseqn20028s:*
7: geneseqn20038s:*
8: geneseqn20038s:*
9: geneseqn20038s:*
10: geneseqn20048s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	324	100.0	324	AAH21111	AAH21111 C. glutam
2	324	100.0	324	AAH65255	AAH65255 C. glutam
3	324	100.0	324	ACA01970	ACA01970 C. glutam
4	324	100.0	1271	AAH21109	AAH21109 C. glutam
5	324	100.0	1271	AAH21112	AAH21112 C. glutam
6	324	100.0	349980	AAH64966	AAH64966 C. glutam
7	34.8	10.7	2000	ADA71938	ADA71938 Rice gene
8	32.8	10.1	3144	ABL27130	ABL27130 Drosophila
9	32.6	10.1	66788	AAH59515	AAH59515 Propionib
10	32.6	10.1	66788	ACF64444	ACF64444 Propionib
11	32.4	10.0	6672	AB168121	AB168121 Ovary can
12	32.4	10.0	6672	ACF39600	ACF39600 Human MEC
13	32.2	9.9	376	ACT20925	ACT20925 DNA clone
14	32.2	9.9	392	ACL20868	ACL20868 DNA clone
15	32.2	9.9	470	ACL20918	ACL20918 DNA clone
16	32.2	9.9	476	ACL20933	ACL20933 DNA clone
17	32.2	9.9	820	AAH05013	AAH05013 Human CDN
18	32.2	9.9	1808	AAH18459	AAH18459 Human CDN
19	31.8	9.8	2091	ADB06513	ADB06513 Altiococ
20	31.8	9.8	2091	ADB06509	ADB06509 Altiococ
21	31.8	9.8	2091	ADB06511	ADB06511 Altiococ
22	31.8	9.8	110000	ADB12064_00	ADB12064 Altiococ
23	31.8	9.8	110000	ADB12064_01	ADB12064 Altiococ

24	31.6	9.8	1404	ADA30052	ADA30052 DNA encod
25	31.4	9.7	2387	AAV52189	AAV52189 Streptoco
26	31.4	9.7	110000	ABH56454_07	ABH56454_07 Continuation (8 of
27	31	9.6	4590	AAH24065	AAH24065 Yeast AOD
28	31	9.6	6415	AAH13212	AAH13212 Enterococ
29	31	9.6	6415	ABH59007	ABH59007 Enterococ
30	30.8	9.5	5949	AAT30053	AAT30053 S. pneumo
31	30.8	9.5	265118	AAH41227	AAH41227 Pyrococu
32	30.6	9.4	789	ACL20863	ACL20863 DNA clone
33	30.6	9.4	789	ACA50939	ACA50939 Prokaryot
34	30.4	9.4	386	ACL20939	ACL20939 DNA clone
35	30.4	9.4	4739	ABX03452	ABX03452 Negative-
36	30.2	9.3	373	ACL20861	ACL20861 DNA clone
37	30.2	9.3	448	ACL20929	ACL20929 DNA clone
38	30.2	9.3	505	ACL20950	ACL20950 DNA clone
39	30.2	9.3	515	ACL20954	ACL20954 DNA clone
40	30.2	9.3	528	ACL20920	ACL20920 DNA clone
41	30.2	9.3	528	ACL20912	ACL20912 DNA clone
42	30.2	9.3	533	ACL20906	ACL20906 DNA clone
43	30.2	9.3	537	ACL20947	ACL20947 DNA clone
44	30.2	9.3	538	ACL20907	ACL20907 DNA clone
45	30.2	9.3	540	ACL20924	ACL20924 DNA clone

ALIGNMENTS

RESULT 1	AAH21111	AAH21111 standard; DNA; 324 BP.
XX	AAH21111	
AC	AAH21111	
XX	AAH21111	
DT	05-SEP-2001 (first entry)	
XX		
DE	C. glutamicum brne DNA.	
XX		
KM	L-amino acid production; brnf; brng; branched-chain amino acid;	
KM	Corynebacterium bacterium; leucine; isoleucine; valine; medicine;	
KM	animal nutrition; ds.	
XX		
OS	Corynebacterium glutamicum.	
XX		
PN	EP1096010-A1.	
XX		
PD	02-MAY-2001.	
XX		
PF	11-OCT-2000; 2000EP-0012057.	
XX		
PR	27-OCT-1999; 99DE-01051708.	
PA	(DEGS) DEGUSSA AG.	
XX	(KERJ) FORSCHUNGSZENTRUM JUELICH GMBH.	
PI	Kennetknecht N, Eggeling L, Sahm H, Pfeifferle W;	
XX	WPI; 2001-391595/42.	
DR	P-PSDB; AAB86248.	
XX		
PT	New export genes from corynebacterium bacteria, useful for increasing	
PT	fermentative production of branched-chain amino acids.	
XX		
PS	Claim 5; Page 16; 23pp; German.	
XX		
CC	This invention describes a novel isolated polynucleotide (I) containing	
CC	at least one sequence that (i) is 70% identical with a sequence that	
CC	encodes at least one of 251 amino acid (aa) (3) or 108 aa (5)	
CC	polypeptides; (ii) encodes a polypeptide at least 70% identical with (3)	
CC	or (5); (iii) is the complement of (i) or (ii); or (iv) contains at least	
CC	15 consecutive bases from (1)-(iii). The invention also describes (a)	
CC	protein derived from sequences of 753 bp (2) or 324 bp (4); (b)	
CC	coryneform microorganisms, especially Corynebacterium, transformed with	
CC	one or more (I), where these are replicative DNA; (c) production of	

[illegible]

Query Match

100.0%; Score 324; DB 6; Length 1271;

Best Local Similarity

100.0%; Pred. No. 5,8e-89;

Matches 324; Conservative

0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGACACGATGATTTCTCTGTAATTCCTGTTGTCGAGATGTCAGTCATTAATTTT 60
 DB 853 ATGACACGATGATTTCTCTGTAATTCCTGTTGTCGAGATGTCAGTCATTAATTTT 912
 QY 61 GCGCTCCGGGCGGCTTCCTGTTAATTCCTTAAGCCCTACGTAATCAATTTGTGGGC 120
 DB 913 GCGCTCCGGGCGGCTTCCTGTTAATTCCTTAAGCCCTACGTAATCAATTTGTGGGC 972
 QY 121 AAAATGGCGATGATGATGTCAGAGATGTCAGATTTTGAACGCGATCAACGTTTGGC 180
 DB 973 AAAATGGCGATGATGATGTCAGAGATGTCAGATTTTGAACGCGATCAACGTTTGGC 1032
 QY 181 AGCAATGGCATGATGATGTCAGAGATGTCAGATTTTGAACGCGATCAACGTTTGGC 240
 DB 1033 AGCAATGGCATGATGATGTCAGAGATGTCAGATTTTGAACGCGATCAACGTTTGGC 1092
 QY 241 GTGGCGCATCTTCTGGCGGTGACGACCTTTGAGCGTTGGCGCTGGACCATGCTT 300
 DB 1093 GTGGCGCATCTTCTGGCGGTGACGACCTTTGAGCGTTGGCGCTGGACCATGCTT 1152
 QY 301 TTGTGGACTGTGATCTTTTC 324
 DB 1153 TTGTGGACTGTGATCTTTTC 1176

RESULT 11

LOCUS

DEFINITION

BD014994 1271 bp DNA linear PART 27-AUG-2002

ACCESSION

VERSION

KEYWORDS

SOURCE

BD014994.1 GI:22555801

ORGANISM

REFERENCE

unclassified.

AUTHORS

TITLE

1 (bases 1 to 1271)

JOURNAL

COMMENT

Kenneth Knecht, N., Sahm, H., Eggeling, L. and Pfeiffer, W.

METHOD

Nucleotide sequence encoding the discharge of branched amino acid, method of isolating the same and utilization thereof

PATENT

Patent: JP 2001169788-A/4

DEPOSIT

OS

Corynebacterium glutamicum ATCC13032

PD

26-JUN-2001

PF

24-OCT-2000

JP 2000324315

PR

27-OCT-1999

DE 19951708.8

PI

NICOLE KENNERKNECHT, HERMANN SAHM, LOTHAR EGGELING, WALTER PI

FFEFERLE

PC

C12N15/09, C07K14/34, C12N1/21, C12P13/06, C12P13/08, C12P13/08, C12N15/09, C12R1/15, C12N1/21, C12R1/15, C12P13/06, C12R1/15, C12N15/00, C12R1/15

CC

brnF

CC

brnE

FH

Key

FT

gene

Location/Qualifiers

(101)..(853)

Location/Qualifiers

(853)..(1176)

Location/Qualifiers

(1176)..(1176)

Location/Qualifiers

(1176)..(1176)

Location/Qualifiers

(1176)..(1176)

Location/Qualifiers

(1176)..(1176)

Location/Qualifiers

(1176)..(1176)

Location/Qualifiers

(1176)..(1176)

Location/Qualifiers

(1176)..(1176)

Location/Qualifiers

(1176)..(1176)

Location/Qualifiers

(1176)..(1176)

Location/Qualifiers

(1176)..(1176)

Location/Qualifiers

(1176)..(1176)

Location/Qualifiers

(1176)..(1176)

Location/Qualifiers

(1176)..(1176)

Location/Qualifiers

(1176)..(1176)

Location/Qualifiers

(1176)..(1176)

RESULT 12

AP454053/c

LOCUS

DEFINITION

AP454053 2105 bp DNA linear BCT 27-JUN-2002

ACCESSION

VERSION

KEYWORDS

SOURCE

AP454053.1 GI:21311379

ORGANISM

REFERENCE

AUTHORS

TITLE

Export of L-isoleucine from Corynebacterium glutamicum: a two-gene-encoded member of a new translocator family

JOURNAL

COMMENT

Submitted (07-NOV-2001) IBT-1, Forschungszentrum, Leo-Brandt Str., Uelrich 52425, Germany

Location/Qualifiers

1..2105

Location/Qualifiers

1..2105

Location/Qualifiers

1..2105

Location/Qualifiers

1..2105

Location/Qualifiers

1..2105

Location/Qualifiers

1..2105

Location/Qualifiers

1..2105

Location/Qualifiers

1..2105

Location/Qualifiers

1..2105

Location/Qualifiers

1..2105

Location/Qualifiers

1..2105

Location/Qualifiers

1..2105

Location/Qualifiers

1..2105

Location/Qualifiers

1..2105

Location/Qualifiers

1..2105

Location/Qualifiers

1..2105

Location/Qualifiers

1..2105

Location/Qualifiers

1..2105

Location/Qualifiers

1..2105

Location/Qualifiers

1..2105

Location/Qualifiers

1..2105

Location/Qualifiers

1..2105

Location/Qualifiers

1..2105

Location/Qualifiers

1..2105

Location/Qualifiers

1..2105

Location/Qualifiers

1..2105

Location/Qualifiers

1..2105

Location/Qualifiers

1..2105

Location/Qualifiers

1..2105

Location/Qualifiers

1..2105

Location/Qualifiers

1..2105

Location/Qualifiers

1..2105

Location/Qualifiers

1..2105

Location/Qualifiers

1..2105

Location/Qualifiers

1..2105


```

RESULT 6
LOCUS AR391953 1271 bp DNA linear PAT 18-DEC-2003
DEFINITION Sequence 1 from patent US 6613545.
ACCESSION AR391953
VERSION AR391953.1 GI:40115724
KEYWORDS
SOURCE
ORGANISM
REFERENCE
  1 (bases 1 to 1271)
  Kemerhnecht, N., Sahm, H., Eggeling, L. and Pfeifferle, W.
  Nucleotide sequences coding for the export of branched amino
  acids, process for the isolation thereof and use thereof
  Patent: US 6613545-A 1 02-SEP-2003;
  Location/Qualifiers
    1..1271
    /organism="unknown"
    /mol_type="genomic DNA"

ORIGIN
Query Match 100.0%; Score 324; DB 6; Length 1271;
Best Local Similarity 100.0%; Pred. No. 5.8e-89;
Matches 324; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGACAACGATGTTCCCTGATTCCTGTTGTCGAGTATGTCATCTACTTT 60
DB 853 ATGACAACGATGTTCCCTGATTCCTGTTGTCGAGTATGTCATCTACTTT 912
QY 61 GCGCTCCGGCGGTTCCGTTCTTAATCCTTAAGCCCCCTACGTAATCACAATTGTGGC 120
DB 913 GCGCTCCGGCGGTTCCGTTCTTAATCCTTAAGCCCCCTACGTAATCACAATTGTGGC 972
QY 121 AAAATGGCGATGTCGATGTCAGAGAAATCCTTGACATTTGACCGCATCAAGTTTGGC 180
DB 973 AAAATGGCGATGTCGATGTCAGAGAAATCCTTGACATTTGACCGCATCAAGTTTGGC 1032
QY 181 AGCAATGCGATGATCTGAAGACTCTAACCCTTGGTCAATGCGCTGCGATTACAGTG 240
DB 1033 AGCAATGCGATGATCTGAAGACTCTAACCCTTGGTCAATGCGCTGCGATTACAGTG 1092
QY 241 GTGGCGCATCTTCTTGCGCGGTGACGACCACTTGTAGGCTTGGCGCTGGACCAATGCT 300
DB 1093 GTGGCGCATCTTCTTGCGCGGTGACGACCACTTGTAGGCTTGGCGCTGGACCAATGCT 1152
QY 301 TTGTTGACTGGTGAATCTTTTC 324
DB 1153 TTGTTGACTGGTGAATCTTTTC 1176

RESULT 7
LOCUS AR391956 1271 bp DNA linear PAT 18-DEC-2003
DEFINITION Sequence 6 from patent US 6613545.
ACCESSION AR391956
VERSION AR391956.1 GI:40115727
KEYWORDS
SOURCE
ORGANISM
REFERENCE
  1 (bases 1 to 1271)
  Kemerhnecht, N., Sahm, H., Eggeling, L. and Pfeifferle, W.
  Nucleotide sequences coding for the export of branched chain amino
  acids, process for the isolation thereof and use thereof
  Patent: US 6613545-A 6 02-SEP-2003;
  Location/Qualifiers
    1..1271
    /organism="unknown"
    /mol_type="genomic DNA"

ORIGIN
Query Match 100.0%; Score 324; DB 6; Length 1271;

```

```

Best Local Similarity 100.0%; Pred. No. 5.8e-89;
Matches 324; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGACAACGATGTTCCCTGATTCCTGTTGTCGAGTATGTCATCTACTTT 60
DB 853 ATGACAACGATGTTCCCTGATTCCTGTTGTCGAGTATGTCATCTACTTT 912
QY 61 GCGCTCCGGCGGTTCCGTTCTTAATCCTTAAGCCCCCTACGTAATCACAATTGTGGC 120
DB 913 GCGCTCCGGCGGTTCCGTTCTTAATCCTTAAGCCCCCTACGTAATCACAATTGTGGC 972
QY 121 AAAATGGCGATGTCGATGTCAGAGAAATCCTTGACATTTGACCGCATCAAGTTTGGC 180
DB 973 AAAATGGCGATGTCGATGTCAGAGAAATCCTTGACATTTGACCGCATCAAGTTTGGC 1032
QY 181 AGCAATGCGATGATCTGAAGACTCTAACCCTTGGTCAATGCGCTGCGATTACAGTG 240
DB 1033 AGCAATGCGATGATCTGAAGACTCTAACCCTTGGTCAATGCGCTGCGATTACAGTG 1092
QY 241 GTGGCGCATCTTCTTGCGCGGTGACGACCACTTGTAGGCTTGGCGCTGGACCAATGCT 300
DB 1093 GTGGCGCATCTTCTTGCGCGGTGACGACCACTTGTAGGCTTGGCGCTGGACCAATGCT 1152
QY 301 TTGTTGACTGGTGAATCTTTTC 324
DB 1153 TTGTTGACTGGTGAATCTTTTC 1176

RESULT 8
LOCUS AX137709 1271 bp DNA linear PAT 30-MAY-2001
DEFINITION Sequence 1 from Patent EP1096010.
ACCESSION AX137709
VERSION AX137709.1 GI:14273886
KEYWORDS
SOURCE
ORGANISM
  Corynebacterium glutamicum
  Corynebacterium glutamicum
  Bacteria; Actinobacteria; Actinobacteriales; Actinomycetales;
  Corynebacteriaceae; Corynebacteriaceae; Corynebacterium.
REFERENCE
  1 Kemerhnecht, N., Eggeling, L., Sahm, H. and Pfeifferle, W.
  Nucleotide sequences coding for branched-chain amino acids export
  Proteins, method for isolating them and their use
  Patent: EP 1096010-A 1 02-MAY-2001;
  Degussa AG (DE) ; FORSCHUNGSZENTRUM JUELICH GMBH (DE)
  Location/Qualifiers
    1..1271
    /organism="Corynebacterium glutamicum"
    /mol_type="unassigned DNA"
    /db_xref="taxon:1718"
    /note="ATCC14752"
    101..856
    /note="unnamed protein product; brnF"
    /codon_start=1
    /transl_table=11
    /protein_id="CAC39923.1"
    /db_xref="GI:14273887"
    /db_xref="REFSEQ:NM_001080492"
    /translation="MKTQEIHSLSLEVPSPKALPEPDKGRYEAIOGKTSIAGL
    GMPIGIAFLGLVIOYGEWMAAPLFGSLFAGSTEMLVIALVGAAPLGLIATTL
    SYWPHVYAPSPFLHVKNPAPRVSVFALIDEAVTAAPAPMSAMRLISQIAFH
    SWSFGGITGTAIAELPFEIKGEPLKCSFYVLTLDSCRTKQIDSLLAGSFTI
    ALVPIPGALPALLIFGLITIRYFLGKAAK"
    853..1179
    /note="unnamed protein product; brnE"
    /codon_start=1
    /transl_table=11
    /protein_id="CAC39924.1"
    /db_xref="GI:14273888"
    /db_xref="REFSEQ:NM_001080493"
    /translation="MTIDPSCILIVAAVAVITPALRAVPFLIKPLRESQFVKAM
    MPAGIATLITASTFRSNAIDKLTITGLLAVAITVAHLIGSRRTLLSVAGITIVF
    GLVNLFF"

```

QY 301 TTGTGGACTGGTGAATCTTTTC 324
DB 301 TTGTGGACTGGTGAATCTTTTC 324

RESULT 4
BD014993
LOCUS BD162491
DEFINITION Novel polynucleotide.
ACCESSION BD162491.1 GI:27868249
KEYWORDS JP 2002191370-A/290.
SOURCE unclassified
ORGANISM unclassified
REFERENCE 1 (bases 1 to 324)
AUTHORS Nakagawa, S., Mizoguchi, H., Ando, S., Hayaishi, M., Ochiai, K., Yokoi, H., Tateishi, N., Senoo, A., Ikeda, M. and Ozaki, A.
TITLE Novel polynucleotide
JOURNAL Patent: JP 2002191370-A 290 09-JUL-2002;
KYOMA HAKKO KOGYO CO LTD
OS Corynebacterium glutamicum
PN JP 2002191370-A/290
PD 09-JUL-2002
PF 15-DEC-2000 JP 2000405096
PI SATOSHI NAKAGAWA, HIROSHI MIZOGUCHI, SEIKO ANDO, MIKIO HAYASHI, KEIICHI OCHIYAI, HARUHIKO YOKOI, NAOKO TATEISHI, AKIHIRO SENOO, MASATO IKEDA, AKIO OZAKI
PC C12N15/09, C07K14/34, C07K16/12, C07K16/40, C12M1/00, PC C12N1/15, C12N1/19, C12N1/21, C12N5/10, C12N9/00, C12N9/02, C12P7/40, C12P13/04, C12P13/08, C12P19/00, C12P19/34, C12P21/02, C12Q1/37, C12Q1/68, G01N33/53, PC G01N33/56,
PC G01N33/569, G01N33/68, G01N37/00//C12P21/08, (C12N1/21, C12R1:15), PC (C12N1/21, C12R1:13), (C12N1/21, C12R1:01), (C12P13/08, C12R1:15), PC C12N15/00, C12N15/00, C12N15/00
CC Novel polynucleotide
FH Key
FT source
FEATURES
source Location/Qualifiers
1..324 /organism="unclassified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"

ORIGIN
Query Match 100.0%; Score 324; DB 6; Length 324;
Best Local Similarity 100.0%; Pred. No. 5e-89;
Matches 324; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGACAACGATTTTCCTCGTATTCCTCTGTTGGCAGTATGTCAGTACTTACTTTT 60
DB 1 ATGACAACGATTTTCCTCGTATTCCTCTGTTGGCAGTATGTCAGTACTTACTTTT 60
QY 61 GGGCTCCGGGCGGTTCCGTTTAACTTAAGCCCTAGCGATGATGAGATTTGGGCG 120
DB 61 GGGCTCCGGGCGGTTCCGTTTAACTTAAGCCCTAGCGATGATGAGATTTGGGCG 120
QY 121 AAAATGGCAGTATGATGTCAGAGACTTTCCTGATTTGACCGCATCAACGTTTGGC 180
DB 121 AAAATGGCAGTATGATGTCAGAGACTTTCCTGATTTGACCGCATCAACGTTTGGC 180
QY 181 AGCAATGCGATGATGTCAGAGACTTTCCTGATTTGACCGCATCAACGTTTGGC 240
DB 181 AGCAATGCGATGATGTCAGAGACTTTCCTGATTTGACCGCATCAACGTTTGGC 240
QY 241 GTGGCGCATCTTCTTGGCGGTGACGACACTTGTGAGCGTTGGCGCATCGTT 300
DB 241 GTGGCGCATCTTCTTGGCGGTGACGACACTTGTGAGCGTTGGCGCATCGTT 300
QY 301 TTGTGGACTGGTGAATCTTTTC 324
DB 301 TTGTGGACTGGTGAATCTTTTC 324

RESULT 5
BD162491
LOCUS BD162491
DEFINITION Novel polynucleotide.
ACCESSION BD162491.1 GI:27868249
KEYWORDS JP 2002191370-A/290.
SOURCE unclassified
ORGANISM unclassified
REFERENCE 1 (bases 1 to 324)
AUTHORS Nakagawa, S., Mizoguchi, H., Ando, S., Hayaishi, M., Ochiai, K., Yokoi, H., Tateishi, N., Senoo, A., Ikeda, M. and Ozaki, A.
TITLE Novel polynucleotide
JOURNAL Patent: JP 2002191370-A 290 09-JUL-2002;
KYOMA HAKKO KOGYO CO LTD
OS Corynebacterium glutamicum
PN JP 2002191370-A/290
PD 09-JUL-2002
PF 15-DEC-2000 JP 2000405096
PI SATOSHI NAKAGAWA, HIROSHI MIZOGUCHI, SEIKO ANDO, MIKIO HAYASHI, KEIICHI OCHIYAI, HARUHIKO YOKOI, NAOKO TATEISHI, AKIHIRO SENOO, MASATO IKEDA, AKIO OZAKI
PC C12N15/09, C07K14/34, C07K16/12, C07K16/40, C12M1/00, PC C12N1/15, C12N1/19, C12N1/21, C12N5/10, C12N9/00, C12N9/02, C12P7/40, C12P13/04, C12P13/08, C12P19/00, C12P19/34, C12P21/02, C12Q1/37, C12Q1/68, G01N33/53, PC G01N33/56,
PC G01N33/569, G01N33/68, G01N37/00//C12P21/08, (C12N1/21, C12R1:15), PC (C12N1/21, C12R1:13), (C12N1/21, C12R1:01), (C12P13/08, C12R1:15), PC C12N15/00, C12N15/00, C12N15/00
CC Novel polynucleotide
FH Key
FT source
FEATURES
source Location/Qualifiers
1..324 /organism="unclassified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"

ORIGIN
Query Match 100.0%; Score 324; DB 6; Length 324;
Best Local Similarity 100.0%; Pred. No. 5e-89;
Matches 324; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGACAACGATTTTCCTCGTATTCCTCTGTTGGCAGTATGTCAGTACTTACTTTT 60
DB 1 ATGACAACGATTTTCCTCGTATTCCTCTGTTGGCAGTATGTCAGTACTTACTTTT 60
QY 61 GGGCTCCGGGCGGTTCCGTTTAACTTAAGCCCTAGCGATGATGAGATTTGGGCG 120
DB 61 GGGCTCCGGGCGGTTCCGTTTAACTTAAGCCCTAGCGATGATGAGATTTGGGCG 120
QY 121 AAAATGGCAGTATGATGTCAGAGACTTTCCTGATTTGACCGCATCAACGTTTGGC 180
DB 121 AAAATGGCAGTATGATGTCAGAGACTTTCCTGATTTGACCGCATCAACGTTTGGC 180
QY 181 AGCAATGCGATGATGTCAGAGACTTTCCTGATTTGACCGCATCAACGTTTGGC 240
DB 181 AGCAATGCGATGATGTCAGAGACTTTCCTGATTTGACCGCATCAACGTTTGGC 240
QY 241 GTGGCGCATCTTCTTGGCGGTGACGACACTTGTGAGCGTTGGCGCATCGTT 300
DB 241 GTGGCGCATCTTCTTGGCGGTGACGACACTTGTGAGCGTTGGCGCATCGTT 300
QY 301 TTGTGGACTGGTGAATCTTTTC 324
DB 301 TTGTGGACTGGTGAATCTTTTC 324

FEATURES	location/Qualifiers
SOURCE	1..324
	/organism="unknown"
	/mol_type="genomic DNA"
ORIGIN	
Query Match	100.0%; Score 324; DB 6; Length 324;
Best Local Similarity	100.0%; Pred. No. 5e-89;
Matches 324; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
QY	1 ATGCAACTGATTTCTCCTGATTTCTCCTGTTGTGCGAGTATGAGTCACTTTC 60
DB	1 ATGCAACTGATTTCTCCTGATTTCTCCTGTTGTGCGAGTATGAGTCACTTTC 60
QY	61 GCGCTCCGGGCGGTTCCGTTCTTAATCCTTAAGCCCCCTACGTGAATCAATTTG 120
DB	61 GCGCTCCGGGCGGTTCCGTTCTTAATCCTTAAGCCCCCTACGTGAATCAATTTG 120
QY	121 AAAATGGGATGTGGATGCGACGAGGATCCTTGATCTTTGACCGCATCAAGCTT 180
DB	121 AAAATGGGATGTGGATGCGACGAGGATCCTTGATCTTTGACCGCATCAAGCTT 180
QY	181 AGCAATGCGATGATCTGAGACTTAACTTTGTGCTCATTCGCGTTCGATTAAGTG 240
DB	181 AGCAATGCGATGATCTGAGACTTAACTTTGTGCTCATTCGCGTTCGATTAAGTG 240
QY	241 GTGGGCGATCTTTCTTGGGGGTGAGACGACCTTTGAGCGTTGGGCGTGGACATG 300
DB	241 GTGGGCGATCTTTCTTGGGGGTGAGACGACCTTTGAGCGTTGGGCGTGGACATG 300
QY	301 TTTGTGAGCTGGTGAATCTTTTC 324
DB	301 TTTGTGAGCTGGTGAATCTTTTC 324
RESULT 2	
AX120374	324 bp DNA linear PAT 11-MAY-2001
LOCUS	AX120374
DEFINITION	Sequence 290 from Patent EP108790.
ACCESSION	AX120374
VERSION	AX120374.1 GI:14037089
KEYWORDS	
SOURCE	
ORGANISM	Corynebacterium glutamicum
REFERENCE	Corynebacterium glutamicum
AUTHORS	Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
TITLE	Corynebacteriaceae; Corynebacterium.
JOURNAL	1 Nakagawa, S., Mizoguchi, H., Ando, S., Hayashi, M., Ochiai, K.,
FEATURES	Yokoi, H., Tateishi, N., Senoh, A., Ikeda, M. and Ozaki, A.
source	Novel polymucleotides
	Patent: EP 108790-A 290-20-JUN-2001;
	KYOWA HAKKO KOGYO CO., LTD. (JP)
	Location/Qualifiers
	1..324
	/organism="Corynebacterium glutamicum"
	/mol_type="unassigned DNA"
	/db_xref="taxon:1718"
ORIGIN	
Query Match	100.0%; Score 324; DB 6; Length 324;
Best Local Similarity	100.0%; Pred. No. 5e-89;
Matches 324; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
QY	1 ATGCAACTGATTTCTCCTGATTTCTCCTGTTGTGCGAGTATGAGTCACTTTC 60
DB	1 ATGCAACTGATTTCTCCTGATTTCTCCTGTTGTGCGAGTATGAGTCACTTTC 60
QY	61 GCGCTCCGGGCGGTTCCGTTCTTAATCCTTAAGCCCCCTACGTGAATCAATTTG 120
DB	61 GCGCTCCGGGCGGTTCCGTTCTTAATCCTTAAGCCCCCTACGTGAATCAATTTG 120
QY	121 AAAATGGGATGTGGATGCGACGAGGATCCTTGATCTTTGACCGCATCAAGCTT 180
DB	121 AAAATGGGATGTGGATGCGACGAGGATCCTTGATCTTTGACCGCATCAAGCTT 180
QY	181 AGCAATGCGATGATCTGAGACTTAACTTTGTGCTCATTCGCGTTCGATTAAGTG 240
DB	181 AGCAATGCGATGATCTGAGACTTAACTTTGTGCTCATTCGCGTTCGATTAAGTG 240
QY	241 GTGGGCGATCTTTCTTGGGGGTGAGACGACCTTTGAGCGTTGGGCGTGGACATG 300
DB	241 GTGGGCGATCTTTCTTGGGGGTGAGACGACCTTTGAGCGTTGGGCGTGGACATG 300
QY	301 TTTGTGAGCTGGTGAATCTTTTC 324
DB	301 TTTGTGAGCTGGTGAATCTTTTC 324

Db	AAATGGCGATGTGGATGCCAGAGAAATCCTTGACATTTTGACCCGATCAAGTTCCG	180
Qy	181 AGCAATGCCATTAATCTGAAGAATCTTAACCTTTGGTCTACTTCCGTTGGATTACAGT	240
Db	181 AGCAATGCCATTAATCTGAAGAATCTTAACCTTTGGTCTACTTCCGTTGGATTACAGT	240
Qy	241 GTGGCGCATTTCTTGAGCGGTGCAGCGACCTTGTGAGCGTTGGCGGTGCACCATCGTT	300
Db	241 GTGGCGCATTTCTTGAGCGGTGCAGCGACCTTGTGAGCGTTGGCGGTGCACCATCGTT	300
Qy	301 TTGTGTGACTGTGAATCTTTTC 324	
Db	301 TTGTGTGACTGTGAATCTTTTC 324	
RESULT 3		
AX137712		
LOCUS	Sequence 4 from Patent EPI096010.	
DEFINITION	AX137712	
VERSION	AX137712.1 GI:14273891	
KEYWORDS		
SOURCE		
ORGANISM	Corynebacterium glutamicum	
REFERENCE	Corynebacterium glutamicum	
AUTHORS	Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;	
TITLE	Corynebacterinease; Corynebacteriaceae; Corynebacterium.	
JOURNAL	1 Kenneknecht, N., Eggeling, L., Sahm, H. and Pfeiffer, W.	
FEATURES	Nucleotide sequences coding for branched-chain amino acids export	
source	proteins, method for isolating them and their use	
	Patent: EP 1096010-A 4 02-MAY-2001;	
	Desussa AG (DE) ; FORSCHUNGSZENTRUM JUELICH GMBH (DE)	
	Location/Qualifiers	
	1..324	
	/organism="Corynebacterium glutamicum"	
	/mol_type="unassigned DNA"	
	/db_xref="taxon:1718"	
	/note="ATCC14752"	
	1..>324	
	/note="unnamed protein product; brne"	
	/codon_start=1	
	/transl_table=11	
	/protein_id="CAC39926.1"	
	/db_xref="GI:14273892"	
	/db_xref="REMBEML:CAC39926"	
	/translation="MTDFSCILIVAVNCVITFALRAVPFLIKPLRESQFVKMAN	
	WMPGILAILTASTFRSNAIDLKTITFGLIAVAITVAHLLGSRRTLLSVAGTIVFV	
	GLVNLF"	
ORIGIN		
Query Match	100.0%; Score 324; DB 6; Length 324;	
Best Local Similarity	100.0%; Prid. No. 5e-89;	
Matches	324; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
Qy	1 ATGAACAATGATTTCTCCTGATATCTCCTTGTGCGAGATATGACGATTAATCTTTT	60
Db	1 ATGAACAATGATTTCTCCTGATATCTCCTTGTGCGAGATATGACGATTAATCTTTT	60
Qy	61 GCGCTCCGGGCGGTTCCGTTCTTAATCCTTAAGCCCTTACGTAATCAAAATTTGGGGC	120
Db	61 GCGCTCCGGGCGGTTCCGTTCTTAATCCTTAAGCCCTTACGTAATCAAAATTTGGGGC	120
Qy	121 AAAATGGGATGTGGATGCCAGAGAAATCCTTGCGATTTGACCGCATCAAGCTTTCG	180
Db	121 AAAATGGGATGTGGATGCCAGAGAAATCCTTGCGATTTGACCGCATCAAGCTTTCG	180
Qy	181 AGCAATGCGATAGATCTGAAGACTTAACTTTGGTCTCATTTGCGGTTCGATTACAGT	240
Db	181 AGCAATGCGATAGATCTGAAGACTTAACTTTGGTCTCATTTGCGGTTCGATTACAGT	240
Qy	241 GTGGCGCATTTCTTGAGCGGTGCAGCGACCTTGTGAGCGTTGGCGGTGCACCATCGTT	300
Db	241 GTGGCGCATTTCTTGAGCGGTGCAGCGACCTTGTGAGCGTTGGCGGTGCACCATCGTT	300

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 19, 2004, 10:57:18 / Search time 1373.24 seconds

(without alignments)
10226.291 Million cell updates/sec

Title: US-10-608-504-6_COPY_853_1176

Perfect score: 324

Sequence: 1 atgcgaactgattctctctctg.....cttgactgctgactcttttc 324

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapept 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : GenEmbl.*
1: gb_da.*
2: gb_hg.*
3: gb_in.*
4: gb_ov.*
5: gb_ov.*
6: gb_ov.*
7: gb_ov.*
8: gb_ov.*
9: gb_ov.*
10: gb_ov.*
11: gb_ov.*
12: gb_ov.*
13: gb_ov.*
14: gb_ov.*
15: gb_ov.*
16: gb_ov.*
17: gb_ov.*
18: gb_ov.*
19: gb_ov.*
20: gb_ov.*
21: gb_ov.*
22: gb_ov.*
23: gb_ov.*
24: gb_ov.*
25: gb_ov.*
26: gb_ov.*
27: gb_ov.*
28: gb_ov.*
29: gb_ov.*
30: gb_ov.*
31: gb_ov.*
32: gb_ov.*
33: gb_ov.*
34: gb_ov.*
35: gb_ov.*
36: gb_ov.*
37: gb_ov.*
38: gb_ov.*
39: gb_ov.*
40: gb_ov.*
41: gb_ov.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	324	100.0	324	6	AR391955
2	324	100.0	324	6	AX120374
3	324	100.0	324	6	AX137712
4	324	100.0	324	6	BD014993
5	324	100.0	324	6	BD162491
6	324	100.0	1271	6	AR391953
7	324	100.0	1271	6	AR391956
8	324	100.0	1271	6	AX137709
9	324	100.0	1271	6	AX137714
10	324	100.0	1271	6	BD014991
11	324	100.0	1271	6	BD014994
12	324	100.0	2105	1	AF454053
13	324	100.0	340000	1	AP005274
14	324	100.0	349980	6	AX120085
15	38.2	11.8	65865	1	AY211269
16	36.6	11.3	4534	1	CFU25133
17	36	11.1	141605	2	AC013732
18	36	11.1	148332	9	AC079586
19	35.8	11.0	2040	8	NTA289862
20	35.6	11.0	137678	8	AP005183
21	35.6	11.0	173270	9	AC024886
22	35.6	11.0	328187	2	AC117393
23	35.4	10.9	113025	2	AC109596
24	35.2	10.9	34494	3	AC006633
25	34.8	10.7	2000	6	AX655353
26	34.8	10.7	86524	4	AC104481
27	34.8	10.7	270748	2	AC121736
28	34.6	10.7	110000	2	AC095209_1
29	34.6	10.7	170475	2	AC136263
30	34.6	10.7	179141	2	AC108890
31	34.6	10.7	224691	2	AC137364
32	34.6	10.7	270850	2	AC095807
33	34.4	10.6	189805	9	AC004169
34	34.4	10.6	190000	2	AC004479
35	34.2	10.6	1686	1	NOCCDH
36	34.2	10.6	12883	1	AE009221
37	34.2	10.6	12909	1	AE008188
38	34.2	10.6	13243	1	AE004602
39	34.2	10.5	150662	2	AC138550
40	34	10.5	167753	9	AC108752
41	34	10.5	168332	2	AC053483
42	34	10.5	169030	2	AC125625
43	34	10.5	251370	5	AL929533
44	33.8	10.4	117645	8	AC144644
45	33.8	10.4	210791	9	CNS01DV7

ALIGNMENTS

RESULT 1
LOCUS AR391955
DEFINITION Sequence 4 from patent US 6613545.
ACCESSION AR391955
VERSION AR391955.1 GI:40115726
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 324)
AUTHORS Kennerknecht, N., Sahm, H., Eggeling, J., and Pfeifferle, W.
TITLE Nucleotide sequences coding for the export of branched chain amino acids, process for the isolation thereof and use thereof
JOURNAL Patent: US 6613545-A 4 02-SEP-2003;

ORIGIN

/plasmid="pBelOBAC11"
/note="end : 17"

Query Match 4.9%; Score 36.8; DB 29; Length 1201;
Best Local Similarity 20.4%; Pred. No. 87;
Matches 73; Conservative 131; Mismatches 153; Indels 1; Gaps 1;

QY 38 TGTGGCCATCCAAAGGCGCCCTGGAACCGATGATPAAAGTTATGCGCCCTACGAATCG 97
DB 823 TTTCCMAMCMCMAMAMMCCCTYCKRAGTANCTCMAMAMMCMSCGSCAAMVVVVVV 882
QY 98 CCGAAGTTAAAAACCTCCCTTGCAGAGTTTGGGCAATGACCGATGATGATGAT 157
DB 883 VVVVVVYCCCCMAGMACCHTATGATGACGMDGASCCGDSVSSGAGCCVMSVGRMYH 942
QY 158 TTGATCTCTTGTATTATCAATACGCTACGAATGATGAGCAGC-CCCATGTTTCCGCG 216
DB 943 YTGAGKBCRGRTTWSGSSGSSGSCGKTGVSAGSSSCSSGSCCSCBCTCGSSSSSS 1002
QY 217 CTGATTTTGGCGGCTCCACCGAAATGCTGATGCTGCTGTTGTGGCCGACGCGCC 276
DB 1003 SSSCBGCTCTYSSSSSSSSSSCKSTYSKGTCTSCYGTGTGKGGCCCGSGCGTC 1062
QY 277 CTGGGCGCCATCGCGCTCACCATTTGCTGTGAATCTCCGACGATTTCTATGCGTTT 336
DB 1063 CBCCCCSCKCGSSGSSSTSKSTGKTBGKSSSKGCGGKSSBSBSBSBNTSSS 1122
QY 337 TCATTCGCCGCTGATGTGTCAAAAACCCCATGCCCCCTTCTATTCGCTTCCGCC 394
DB 1123 TTTTBSSSSTSTTSBSGSBTTBSBSSTSTSBTYTSTTBTSSSBSTSSSKB 1180

RESULT 15
BE013856/c 539 bp mRNA linear EST 09-JUL-2000
LOCUS BE013856
DEFINITION 125354 MARC 1P1G Sus scrofa cDNA 5', mRNA sequence.
ACCESSION BE013856
VERSION BE013856.1 GI:8274843
KEYWORDS EST.
SOURCE Sus scrofa (pig)
ORGANISM Sus scrofa
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.
1 (bases 1 to 539)
Fahrenkrug,S.C., Smith,T.P.L., Freking,B.A., Cho,J., White,J.,
Vallet,J., Wise,T., Rohrer,G.A., Perlea,G., Sultana,R.,
Quackenbush,J. and Keefe,J.W.
Porcine gene discovery by normalized cDNA-library sequencing and
EST cluster assembly
Mamm. Genome 13 (8), 475-478 (2002)

TITLE
JOURNAL
MEDLINE
PUBMED
COMMENT

Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called and alt trimmed with phred
v0.980904.e. Vector identified by cross_match with the -minscore 18
and -mismatch 12 options.
PCR Primers
FORWARD: AGGAACAGCTATGACCAT
BACKWARD: GTTTCCAGTCACGACG
Plate: 52 row: G column: 21
Seq primer: ATTAGGTGACACTATG.
Location/Qualifiers
1..539

FEATURES

source
/organism="Sus scrofa"
/mol_type="mRNA"
/db_xref="taxon:9823"
/feature_type="pooled"

ORIGIN

Query Match 4.9%; Score 36.6; DB 10; Length 539;
Best Local Similarity 54.0%; Pred. No. 67;
Matches 75; Conservative 0; Mismatches 64; Indels 0; Gaps 0;

QY 211 TCGGCTGATTTTGGGCGCTCCACGAATGCTGTATCGCCCTGTGTGGCGCA 270
DB 445 TGTGCTGGGCTTGGCCAGGCGCAGGTGAGGGGTGTGAGGCTGTACAGATCA 386
QY 271 GCGCCCTGGGCGCATCGGCTACCACTTCTGTGAATTCGCCACGATTTCTAT 330
DB 385 CCGGCCGAGGAGACAGCCACCCGCCACCTCTGACAGCCAGGAGCGCATCAAG 326
QY 331 GCGTTTCATTCCCGCTGC 349
DB 325 GTGTCTGCAAGCCGCTGC 307

Search completed: April 17, 2004, 20:01:49
Job time: 2438.38 secs

/lab_host="DH10B"
/clone_lib="MARC 1P1G"
/note="vector: PCMV SPORTE; Site 1: NotI, Site 2: SalI;
Library made from pooled tissue from day 11, 13, 15, 20,
and 30 embryos."

Library was constructed by Life Technologies, a division of Invitrogen. Contact : Feng Liang Email : fliang@life-tech.com URL : <http://fulllength.invitrogen.com/> Invitrogen Corporation 1600 Faraday Avenue Genoscope sequence ID : CSODI074BBI20P1.

FEATURES

source

Location/Qualifiers

```
1..1108
/organism="Homo sapiens"
/mol_type="RNA"
/db_xref="taxon:9606"
/clone="CSODI074YC24"
/cisue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime end enriched, double-strand cDNA was digested with NotI and cloned into the NotI and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."
```

ORIGIN

Query Match 5.0%; Score 37.4; DB 13; Length 1108;
Best Local Similarity 16.3%; Pred. No. 57;
Matches 65; Conservative 145; Mismatches 189; Indels 0; Gaps 0;

```
114 CTCCTGCTGAGGTTGGGCGATGACCCGATGATGCTTGGCTTGGTAT 173
174 TCAATGCGCTAGGAGGAGGCGAGCCGACGCTGTTCCGGCTGATTTCCGGGCTC 233
480 TSTTTCGTCGCGCTTCTKTDCTTKKCKSMCKWCCSSTKCTCKSKKSKTKT 421
234 CACCGAATTCGTGATCGCCCTGTTGGGCGAGCGCCCTGGGCGCATGCGCT 293
420 WCKKTTTTKKAKCKACTKCKSKTKWCKSKCKSKCKSKCKSKCKSKCKSKCK 361
294 CACCACTTCGTGATCGCCGACGATCTGAGGCTTCAATCCGCTGATCT 353
360 SSACKSKTKCKCTKCKSKSKKTKKSKKCKSKCKSKCKSKCKSKSKSKTKST 301
354 GATCAAAACCCATGCGCGTTCATTTGCGGTTTTCGCGCTTATGACGAACCTAC 413
300 TSTATTATATACAMTSSSSSMKMTGTTMKKSDTSSMKASCKCTKSTSTTS 241
414 AGTACTGCGGCGAGCGCCGAGCTGTGGGCGGCGGACCTATCATGAATGAC 473
240 KSSATATGKSTCTSMCKSSCSCKSSSTGSCCKSGMTSTTTSTKSTGSCCK 181
474 GTTCACCTCCTACTGCGTATTCGGCGGCTCTACCGGAGT 512
180 TSMTKMTKTSSTSTSKKMKKAKKSTSATSSKCK 142
```

RESULT 11
BG845919/c 730 bp mRNA linear EST 29-MAY-2001
LOCUS 1024011F06.X1 C. reinhardtii CC-1690, normalized, Lambda Zap II
DEFINITION Chlamydomonas reinhardtii cDNA, mRNA sequence.

ACCESSION BG845919
VERSION BG845919.1 GI:14227103
KEYWORDS EST
SOURCE Chlamydomonas reinhardtii
ORGANISM Chlamydomonas reinhardtii
Chlamydomonas reinhardtii
Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
Chlamydomonadales; Chlamydomonadales.

REFERENCE Grossman, A., Davies, J., Federpspiel, N., Harris, E., Lefebvre, P.,
AUTHORS McErmott, J.P., Sillitow, C., Stern, D. and Strzycki, R.
TITLE Analyses of the Chlamydomonas reinhardtii Genome: A Model,
Unicellular System for Analyzing Gene Function and Regulation in
Vascular Plants; project phase 2
COMMENT Unpublished (2000)
CONTACT: Charles Hauser
DCMB Box 91000
Duke University

Durham, NC 27708-1000
Tel: 919 613 8159
Fax: 919 613 8177
Email: chause@duke.edu

FEATURES

source

Location/Qualifiers

```
1..730
/organism="Chlamydomonas reinhardtii"
/mol_type="RNA"
/strain="CC-1690 wild type mt+ 21gr"
/db_xref="taxon:3055"
/clone_lib="C. reinhardtii CC-1690, normalized, Lambda Zap II"
/notes="vector: pBluescript II SK-; site 1: EcoRI, site 2: XhoI; This library, constructed by John Davies and Jeffrey McErmott, combines cDNAs from CC-1690 cells grown to mid-log phase in TAP (acetate-containing) medium in the light, TAP medium in the dark, HS (minimal) medium in ambient levels of CO2 and HS medium bubbled with 5% CO2. PolyA mRNA was purified from each sample, pooled and cDNA synthesized. The cDNA was directionally cloned into lambda Zap II (Stratagene) in the EcoRI (5') and XhoI (3') sites. pBluescript II SK- plasmids were excised from the lambda Zap clones by superinfection with Exsist (Stratagene) phage. The library was normalized using method 4 described in Bonaldo et al (1996) Genome Research 6: 791-806."
```

ORIGIN

Query Match 4.9%; Score 37.2; DB 12; Length 730;
Best Local Similarity 49.5%; Pred. No. 53;
Matches 96; Conservative 0; Mismatches 98; Indels 0; Gaps 0;

```
202 CCACTGTTTCCGCGGCTTTGCGGCGCTCCACGGAATGCTGCTGCTGCTT 261
691 CCGCGGTCAAGGCGGCGGAGCGCCGCCCGCCCAAGTTAGAGGATCGCGCGC 632
262 GTGGCGGAGCGCCCTGCGGCGGCGGCGCTCACACATGCTGTGAACTTCCG 321
631 TTGTGGCGCGGCGGCTGACCGCCGCTGATCAGAGGTGGTGCAGTGTCCAG 572
322 GTATTCTATGCTTTCAATTCCTCCCTGATGCTGATCAAAACCCATGCG 381
571 GACGCGATGACATGCTGCTCCCGCTGCGCTGAGGAGGCGCTCATGTTGG 512
382 TCGGTTTTCGCGCT 395
511 TCGGCGCGCGCGGT 498
```

RESULT 12
BF101158 1454 bp mRNA linear EST 19-OCT-2000
LOCUS 60175473P1.NCI_CGAP_Mam1 Mus musculus cDNA clone IMAGE:3983528 5',
DEFINITION mRNA sequence.

ACCESSION BF101158
VERSION BF101158.1 GI:10883684
KEYWORDS EST
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE NIH-MGC <http://mgi.nci.nih.gov/>
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1995)
CONTACT: Robert Strausberg, Ph.D.
Email: cgabs-r@mail.nih.gov

COMMENT Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNI)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILNI at:
<http://image.llnl.gov>

REFERENCE 1 (bases 1 to 697)
AUTHORS Myler, P. J., Vogt, C., Menden, H., Robertson, L., Sisk, E.,
Fazelinia, G., Aggarwal, G., Nelson, S., Seyler, A., Worthey, E.,
Stuart, K. and Ragland, M.
TITLE Leishmania major Friedlin BAC End Sequences
JOURNAL Unpublished (2002)
COMMENT Other GSSs: LB01712a.d.T7.1
Contact: Myler PJ
Seattle Biomedical Research Institute
4 Nickerson Street, Seattle, WA 98109-1651, USA
Tel: 206 284-8846
Fax: 206 284-0313
Email: mylerpj@sebrt.org
Seq primer: SP6
Class: BAC ends.

FEATURES
Source Location/Qualifiers
1..697
/organism="Leishmania major"
/mol_type="Genomic DNA"
/strain="Friedlin"
/db_xref="taxon:5664"
/clone="LB01712a"
/lab_host="E. coli GeneHogs + TrfA"
/clone_lib="Leishmania major Friedlin BAC Library"
/note="Vector: pCG270; Site 1: HindIII; Genomic DNA from
Leishmania major Friedlin in agarose blocks was partially
digested with HindIII, size selected, and ligated with
HindIII-digested pCG270 vector DNA. 10368 clones were
picked and arrayed in 384- and 96-well plates. Library
construction and arraying was carried out by ResGen
Corporation and clones and filters are available from
them"

ORIGIN
Query Match 5.0%; Score 37.4; DB 28; Length 697;
Best Local Similarity 52.2%; Pred. No. 46;
Matches 83; Conservative 0; Mismatches 76; Indels 0; Gaps 0;

Db 181 GGCTACGAATGTTGGGAGCCCACTGTTTCGGGCTGATTTTCGGGGCTCCACCGAA 240
579 GGAGCCGATCGTGTGTCGCGACCTCCGACCCACATCGCTCTGCGCGCGCG 520
Dy 241 ATGCTGTCATCGCCCTCTGTTGTGGGGGAGCGCCCTGGGGCGCATGCGCTCACCA 300
Dy 519 CACGAGTCGGCGCGCGCTGCTGTCAGCCTCTGTCGCGCTTGGCGGACCCGTTACCCAG 460
Dy 301 TTGCTGGGAACTTCGCGCACGATATTCATGCGTTTCA 339
Dy 459 CCACTGTCACGTGTGTCATGCTGTCGTCGTCACATCTCA 421

RESULT 10
BX382084/c BX382084.c
LOCUS 1108 bp RNA linear EST 08-MAY-2003
DEFINITION BX382084 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens CDNA
clone CS001074Yc24 5-PRIME, mRNA sequence.
ACCESSION BX382084
VERSION BX382084.1 GI:30455015
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidi; Homo.
1 (bases 1 to 1108)
Li, W.B., Gruber, C., Jessee, J. and Playates, P.
Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqret@genoscope.cns.fr, Web : www.genoscope.cns.fr

Db	1100	SCNMHBSGCKGCCKKMKCMKMCBCECMCBKVCYVXXXXXCKKKKKCCXXCCNNKK	1159
Qy	599 GC	600 : 	
Db	1160 BC	1161	
RESULT 6			
CNS0091P			
LOCUS			
DEFINITION	CNS0091P	925 bp	DNA linear GSS 03-JUN-1999
	Drosophila melanogaster genome survey sequence TET3 end of BAC #		
	BACR19616 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.		
ACCESSION	AL053013		
VERSION	AL053013.1	GI:4934461	
KEYWORDS	GSS.		
SOURCE	Drosophila melanogaster (fruit fly)		
ORGANISM	Drosophila melanogaster		
	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila. 1 (bases 1 to 925)		
REFERENCE	Genoscope.		
AUTHORS	Direct Submission		
TITLE	Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage : D. Nat. etoc envy adex + PRIME (F-mail : sequef@genosco.cns.fr		
JOURNAL			

COMMENT

The BDGP is constructing a physical map of the *Drosophila melanogaster* genome using these BACs. For further information please see <http://www.fruitfly.org> The BDGP *Drosophila melanogaster* BAC library was prepared by Kaetzo Osoegawa and Aaron Mammoser in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPci-98 and was constructed by partial *EcoRI* digestion of *Drosophila* DNA provided by the BDGP from the isogenic strain y2; cn bw sp, the same strain used for the BDGP's p1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

```

FEATURES
    source
        1..925
            location/Qualifiers
            /organism="Drosophila melanogaster"
            /mol_type="genomic DNA"
            /db_xref="taxon:7227"
            /clone="BACR19p16"
            /clone_1b="Rc1-98"
            /note="end : TET3"

```

ORIGIN	
Query Match	5.2%
Best Local Similarity	14.2%
Best Match	37
Conservative	117
Mismatches	106
Indels	
Score	39
DB	29
length	925

	Query Match	5.2%	Score 39;	DB 29;	Length 925;
	Best Local Similarity	14.2%;	Pred No. 19;		
	Matches	Similarity	Conservative	Mismatches	Indels
				Gaps	
Qy	195 GCGAGCCCATCTGTTTCCGGCCTGAATTTCCGGGCCTCACCGAAATGCTGTCAATCCG	254			
Db	638 SSSSSSTSSSTSSTSKSSSSSGSSSSSSSYTTSKSTASGGSGMSAAGSGSSTGSTSS	697			
Qy	255 CCTCGTTGGGGGCGACGCCCTTCGGGCGCATCGGCGTCAACACATTGGTGTAACCT	314			
Db	698 SSSSSSTSSSSSVSSGSKSSTSSSSSBSSSGSSSSSSTSSBSCTSTSSSSSVSSST	757			
Qy	315 CGCGCACGATTCATGTCGTTTCATTCATCCGCTGACATGTCGTCAAAACCCCATGGCCG	374			
Db	758 CSCCTCCCSYSSSTSSSSSTSWGSTGSSSSSVSSSDSTSTCCSCCTTCCTCCT	817			
Qy	375 TTTCATATGGGTTTGGCGGCTTATCGACGAAGCTACGACATCGGCGCACAGCCGC	434			
Db	818 YEMBCYTSTSCGSSSSSGKGVTCKCGCGCGSSSTMGHGTISACGSSSSSCSSSVVS	877			

```

QY      435 AGGCTGCTCGGCGTGCGAC 454
      :: : :::: |::|::
Db      878 SSKSASASSSVSSSGSGVS 897

```

RESULT	7
BX361080	
LOCUS	
DEFINITION	BX361080 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens CDNA clone CS01079YN16 3-PRIME, mRNA sequence.
ACCSSION	BX361080
VERSION	BX361080
KEYWORDS	
SOURCE	EST.
ORIGIN	Homo sapiens (human)
CDS	Homo sapiens

REFERENCE	1 (bases 1 to 1201)
AUTHORS	Li, W.B., Gruber, C., Jesssee, J. and Polayes, D.
TITLE	Full-length cDNA libraries and normalization
JOURNAL	Unpublished (2001)
COMMENT	Contact: Genoscope
	Genoscope, Centre National de Séquençage

Genoscope - Centre National de Séquençage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr Web: www.genoscope.cns.fr
Library was constructed by life technologies, a division of
Invitrogen Contact: Feng liang Email: fliang@lifetech.com URL:
<http://fulllength.invitrogen.com/> Invitrogen Corporation 1600
Paradise Avenue genoscope séquence ID : CS0D1079DC08NP1.

FEATURES	Location/Qualifiers
source	1. .1201

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSOD107YIN16"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

```

ORIGIN

Query Match	5.1%;	Score 38.6;	DB 13;	length 1201;
Best Local Similarity	10.5%;	Pred: NO. 28;		
Matches 34;	Conservative 147;	Mismatches 142;	Indels 0;	Gaps 0

19 ATTCATTCAAGCCCTGGAGGTCTCCCATCCAAAGCAGCCCTGAAACCATGATTAAGG

Ddb

8 25 MDDMMNNMKNNDKKKKNDEKKNMKAKAKKKKKKKKKD KDKKNMDKD KKNMKGKKKAMKGNAMA KK

Ddb

885 KKKKKDAAKKKKKMAKKKKGLAALAKKKKKKKKKKKKKNNNNHKKCKMKCKMKEKKKKCDMMHKM

139 TACCGATTGTAATGCGTTGCTCTCTTGCTATATCAACGGCTACGATGTTGGGC

Db 945 KKKKKKKKT KKKTTTTTWTTTTTTHTMTMMHTTTTTTTTMTTTTTTHNKKKKKKKKKKK

Ddb

1005 KNDNNTMMMKKKQMHKMKKEKKKKKKNNMKKKQNKEMNKKNMNMKMKKKKKNNMTNKIK

259 GTTGTGGGGCGCAGCGCCCCCTGGGCGGCATCGCGCTACCAATGTGCTGTGAATTCCG

Db 1065 KKKKKKKCCCCCBKBCCCCCCKYKCKBKRYCCYKKKKKKKKKKYCCCBCCCKKKCBVC

[illegible]

RESULT 8

RESULT 8

OY		491	TATTCGGGCTCACCAGGAATGGCAATCGAAGATTGATTTCTTGTGAATTAAAGGCC	550
Dd		535	GTTTCGCGCGCCGCACCCGATTCACACTCGGCTTCGCGAGCGCTCTTAAGATTCAAGTC	594
OY		551	TCGAGTTCCGCCCCCTTGCTCTCTCTCTTGTGTACGC	583
Dd		595	ACGAGAGGCGCCTTGTCTCTCTCTCCCTGCCCCG	627
 RESULT 4 CONS009WA/c				
LOCUS				
DEFINITION				
Drosophila melanogaster genome survey sequence T7 end of BAC #				
BACR20K04 of RPCL-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.				
AL054013 GI:4935590				
VERSION				
KEYWORDS				
SOURCE				
ORGANISM				
Drosophila melanogaster (fruit fly)				
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;				
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;				
Ephydroidea; Drosophilidae; Drosophila.				
1 (bases 1 to 787)				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
Direct Submission				
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage ;				
BP 191 91006 EVRY cedex - FRANCE (E-mail : sequef@genoscope.cns.fr				
Web : www.genoscope.cns.fr)				
Determination of this BAC-end sequence was carried out as part of a				
collaboration with the Berkeley Drosophila Genome Project (BDGP).				
The BDGP is constructing a physical map of the Drosophila				
melanogaster genome using these BACs. For further information				
please see http://www.fruitfly.org The BDGP Drosophila				
melanogaster BAC library was prepared by Kazuo Oosawa and				
Aaron Mamoser in Peter de Jong's laboratory in the Department of				
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,				
NY. The library is named RPCL-98 and was constructed by partial				
EcoRI digestion of Drosophila DNA provided by the BDGP from the				
isogenic strain y2; cn bw sp, the same strain used for the BDGP's				
p1 and EST libraries. A more detailed description of the library				
and how to order individual BAC clones, the entire library, or				
filters for hybridization from the BACPAC Resource Center can be				
found at http://becpac.med.buffalo.edu/drosophila_bac.htm.				
Location/Qualifiers				
1..787				
/organism="Drosophila melanogaster"				
/mol_type="genomic DNA"				
/db_xref="taxon:7227"				
/clone_lib="BACR20K04"				
/clone_1lb="RPCL-98"				
/note="end : T7"				
 ORIGIN				
Query Match				
Best Local Similarity 45.5%; Pred. No. 14;				
Matches 76; Conservative 15; Mismatches 76; Indels 0; Gaps 0				
OY		552	CGAGTTGCGCCCTTGCTCTCTCTTGTACAGCTGATTTGGATTCTCTGCCGAACAAGAAA	611
Dd		334	CTATTCCTGCTCTTCTGCTCTTGTATTAAMKKMCTGGAAATCCCKTRVAGARPAAC	275
OY		612	GCAGATCCCTCTCTGTGTCTGACAGATTGAGCTCACCATTCGCTTGTGGTAATCC	671
Dd		274	AATTGTTKTTTTCVACWGRCTTAAGGATAATTTCRTAATTTTTTTTGTATTT	215
OY		672	AGGTGAGCCCTATTGCGGCGCTGCTGATCTTCTTGAGTCTGTGA	718
Dd		214	TAAATTGCCATATGATTATATCGAATTTTATTATGTCGCGTA	168
 RESULT 5 BX38196L				

LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT	FEATURES
EX381961	1201 bp	NRNA	linear	EST 08-MAY-2003								
EX381961	Homo sapiens PLACENTA COT 25-NORMALIZED	clone CS0D1072YF05	3-PRIME, mRNA	sequence.								
EX381961	EST.	GI:30453007										
Homo sapiens (human)												
Homo sapiens												
Eukaryote, Metazoa, Chordata, Craniata, Vertebrata, Euteleostomi, Mammalia, Eutheria, Primates, Catarrhini, Homnidae, Homo.												
1 (bases 1 to 1201)												
L1, W.B., Gruber, C., Jesse, J. and Polayes, D.												
Full-length cDNA libraries and normalization												
Unpublished (2001)												
Contact: Genoscope												
Genoscope - Centre National de Sequencage												
BP 191 91006 Evry cedex - France												
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr												
Library was constructed by life technologies, a division of invitrogen. Contact : Feng liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/Invitroden Corporation 1600												
Faraday Avenue Genoscope sequence ID : CS0D1072CC03NP1.												
Location/Qualifiers												
1..1201												
/organism="Homo sapiens"												
/mol_type="mRNA"												
/db_xref="taxon:9606"												
/clone="CS0D1072YF05"												
/issue_type="PLACENTA COT 25-NORMALIZED"												
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"												
/note="1st strand cDNA was primed with a NotI-oligo (dt) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pcWVSPORT 6 vector. Library was normalized."												
ORIGIN												
Query Match	5.2%; Score 39.2; DB 13; Length 1201;											
Beef Local Similarity 6.3%; Pred. No. 19;												
Matches 34; Conservative 210; Mismatches 288; Indels 0; Gaps 0;												
59 TGAACGATGATAAAGTTATGCGGCGCTACGAATCGCGCAAGTCTAAAACTCCC	118											
620 KKK	679											
119 TTGCTGCGAGTTTGGCGCATGACCGATGTTGATTTGGTCTGTGTTATCAAT	178											
660 KKK	739											
179 ACGGCTACGAATGTGTGGCAGCCCCACTGTTTCCGGCTGATTTTGGCGGACTCC	238											
740 MNNNNCNCKKK	799											
239 AAATGCTGTCATGCGCCTCGTGTGGCGGAGCGCCCTGGGCGCATCGCGCTCACCA	298											
800 MCCCKKK	859											
299 CATTGCTGGTGAATCTCCGACAGTATTTCTATGCGTTTCAATTCGCGTGCATGTGTC	358											
860 MNNNNKK	919											
359 AAAACCCCATTTCCCGTTTCTATTGGGTTTGGCGCTATGAGAGAACCTACGAGTCA	418											
920 MNNCKKK	979											
419 CTGCGCGCAGGCGCCGACAGTGTGCGCGGTGCGGCTTATTCAT												

Page 2

Query Match	5.2%	Score 39.4	DB 29	Length 743
Best Local Similarity	53.6%	Pred. No. 13		
Matches	82	Conservative	0	Mismatches 71; Indels 0; Gaps 0.

Qy	431	CCCCAGGCGGCGCGCGTTCGACATTATCTCAATGCAAAATAGCGTTTCACTCTCTACGCG	490
Ds	475	CTCTCCGCTCTCCGCGCGCTTCTCTCTCTCCCACTCTCTCATCTCTCCGATTGCGAAATCTTCGG	534

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 17, 2004, 06:48:18 ; Search time 2438.38 Seconds
(without alignments)
9221.792 Million cell updates/sec

Title: US-10-608-504-6_COPY_101_853

Perfect score: 753
Sequence: 1 gtgcacaaacgcagagat.....tcttggaagagctgctaa 753

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: em_eebha:*
2: em_eebha:*
3: em_eebha:*
4: em_eebha:*
5: em_eebha:*
6: em_eebha:*
7: em_eebha:*
8: em_eebha:*
9: gb_eeb1:*
10: gb_eeb2:*
11: gb_eeb3:*
12: gb_eeb4:*
13: gb_eeb5:*
14: gb_eeb6:*
15: em_eeb7:*
16: em_eeb8:*
17: em_eeb9:*
18: em_eeb10:*
19: em_eeb11:*
20: em_eeb12:*
21: em_eeb13:*
22: em_eeb14:*
23: em_eeb15:*
24: em_eeb16:*
25: em_eeb17:*
26: em_eeb18:*
27: em_eeb19:*
28: gb_eeb20:*
29: gb_eeb21:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	54.6	7.3	473	28	AQ936633 HSU41-905
2	45	6.0	1201	13	BX356664 BX356664
3	39.4	5.2	743	29	CC684922 CC684922
4	39.4	5.2	787	29	CNS009WA AL054013 Drosophila

5	39.2	5.2	1201	13	BX381961	BX381961
6	39	5.2	925	29	CNS0091P	AL055013 Drosophila
7	38.6	5.1	1201	13	BX361080	BX361080
8	38	5.0	835	14	CB657749	CB657749
9	37.4	5.0	697	28	BH887860	BH887860
10	37.4	5.0	1108	13	BX382084	BX382084
11	37.2	4.9	730	12	BG845919	BG845919
12	37.2	4.9	1454	10	BP101158	BP101158
13	36.8	4.9	464	28	BH805539	BH805539
14	36.8	4.9	1201	29	CNS016BR	AL056545 Drosophila
15	36.6	4.9	539	10	BE013856	BE013856
16	36.6	4.9	942	29	CNS000CT2	AL059804 Drosophila
17	36.4	4.8	298	12	BE525053	BE525053
18	36.2	4.8	673	12	B0287010	B0287010
19	36.2	4.8	842	13	BX331708	BX331708
20	36.2	4.8	902	29	CQ358991	CQ358991
21	36.2	4.8	1101	13	B0678281	B0678281
22	36.2	4.8	559	13	BX384405	BX384405
23	36	4.8	666	28	CC170529	CC170529
24	36	4.8	784	29	CG250542	CG250542
25	36	4.8	836	29	CG449584	CG449584
26	36	4.8	852	28	B2644270	B2644270
27	36	4.8	863	29	CG449578	CG449578
28	35.8	4.8	410	9	AI138160	AI138160
29	35.8	4.8	641	28	B2375719	B2375719
30	35.8	4.8	641	29	CG668174	CG668174
31	35.8	4.8	696	29	CC654573	CC654573
32	35.8	4.8	980	29	CG326716	CG326716
33	35.6	4.7	327	10	BF543765	BF543765
34	35.6	4.7	515	13	BX424977	BX424977
35	35.6	4.7	592	9	AI293892	AI293892
36	35.6	4.7	657	13	CA109167	CA109167
37	35.6	4.7	683	14	CP850119	CP850119
38	35.6	4.7	684	28	CC337836	CC337836
39	35.6	4.7	798	12	BG855128	BG855128
40	35.6	4.7	900	29	CG371461	CG371461
41	35.6	4.7	977	29	CNS000X7	AL076880 Drosophila
42	35.6	4.7	985	13	BX359916	BX359916
43	35.6	4.7	1201	13	BX381971	BX381971
44	35.6	4.7	1398	14	CP65113	CP65113
45	35.4	4.7	454	9	AI575130	AI575130

ALIGNMENTS

RESULT 1
LOCUS HSU41-905 Human Ncti clones Homo sapiens genomic, genomic survey
DEFINITION
ACCESSION AQ936633 473 bp DNA linear GSS 23-AUG-2000
VERSION AQ936633
KEYWORDS
SOURCE GSS. sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
Zabarovskiy E.R., Gazatullin R., Podowski R.M., Zabarovskiy V.V.,
1 (bases 1 to 473)
Li J., Protchenko O.V., Kozlov S., Petrenko L., Skobeleva N.,
Li J., Protchenko O.V., Kaashub V., Ernerberg I., Winberg G. and
Mahlstedt C.
Ncti clones in the analysis of the human genome
Nucleic Acids Res. 28 (7), 1635-1639 (2000)
JOURNAL MEDLINE
PUBMED
COMMENT
Contact: Podowski R.M.
Center for Genomics Research
Karolinska Institute
17177 Stockholm, Sweden
Tel: +46-8-728-6372
Fax: +46-8-337983

Tue Apr 20 06:47:25 2004

us-10-608-504-6_copy_101_853.rnpn

Search completed: April 17, 2004, 20:13:05
Job time : 88.0151 secs

Tue Apr 20 06:47:25 2004

us-10-608-504-6_copy_101_853.rnpn

Page 5

RESULT 12

US-60-548-091-19934
; Sequence 19934, Application US/60548091
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: STROKE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001506
; CURRENT APPLICATION NUMBER: US/60/548,091
; CURRENT FILING DATE: 2004-02-27
; NUMBER OF SEQ ID NOS: 24433
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19934
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-60-548-091-19934

Query Match 4.4%; Score 33; DB 7; Length 201;
Best Local Similarity 52.7%; Pred. No. 0.67;
Matches 69; Conservative 1; Mismatches 61; Indels 0; Gaps 0;

QY 242 TCGTGTATCGCCCTCGTGTGTGGCGAGCGCCCTCGGCGCCATCGCGCTACACCAT 301
DB 46 TCGTGTATCGCCCTCGTGTGTGGCGAGCTCGCTCGTGTGTGTGTGTGTGTGTGT 105
QY 302 TCGTGTATCGCCCTCGTGTGTGGCGAGCTCGCTCGTGTGTGTGTGTGTGTGTGT 361
DB 106 GCTTACAGGACACCGCGCAATTTCATCCATGAGCGCTTCCGCTGTGGATTCTACC 165
QY 362 ACCCATGCGC 372
DB 166 TGCCTCTGACC 176

RESULT 13

US-60-548-091-19938
; Sequence 19938, Application US/60548091
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: STROKE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001506
; CURRENT APPLICATION NUMBER: US/60/548,091
; CURRENT FILING DATE: 2004-02-27
; NUMBER OF SEQ ID NOS: 24433
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19938
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-60-548-091-19938

Query Match 4.4%; Score 33; DB 7; Length 201;
Best Local Similarity 52.7%; Pred. No. 0.67;
Matches 69; Conservative 1; Mismatches 61; Indels 0; Gaps 0;

QY 242 TCGTGTATCGCCCTCGTGTGTGGCGAGCGCCCTCGGCGCCATCGCGCTACACCAT 301
DB 49 TCGTGTATCGCCCTCGTGTGTGGCGAGCTCGCTCGTGTGTGTGTGTGTGTGTGT 108
QY 302 TCGTGTATCGCCCTCGTGTGTGGCGAGCTCGCTCGTGTGTGTGTGTGTGTGTGT 361
DB 109 GCTTACAGGACACCGCGCAATTTCATCCATGAGCGCTTCCGCTGTGGATTCTACC 168
QY 362 ACCCATGCGC 372
DB 169 TGCCTCTGACC 179

RESULT 14
US-10-767-701-9054

Sequence 9054, Application US/10767701

; GENERAL INFORMATION:
; APPLICANT: KOVALIC, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 9054
; LENGTH: 1028
; TYPE: DNA
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-CLUS73927_1
US-10-767-701-9054

Query Match 4.4%; Score 32.8; DB 6; Length 1028;
Best Local Similarity 54.0%; Pred. No. 1.7;
Matches 67; Conservative 0; Mismatches 57; Indels 0; Gaps 0;

QY 609 AAGCAGATCCCTTCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 668
DB 746 ATAGCCCTTCTCTGT 805
QY 669 TCCAGGTACAGCCCTTATTTTGGCGCGCTGATCTTCTGTGTGTGTGTGTGTGTGT 728
DB 806 GCCCAATACAGCCCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 865
QY 729 CTTC 732
DB 866 GTCC 869

RESULT 15

US-10-767-701-23189
; Sequence 23189, Application US/10767701
; GENERAL INFORMATION:
; APPLICANT: KOVALIC, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 23189
; LENGTH: 519
; TYPE: DNA
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: 14593439
US-10-767-701-23189

Query Match 4.3%; Score 32.6; DB 6; Length 519;
Best Local Similarity 52.6%; Pred. No. 1.4;
Matches 71; Conservative 0; Mismatches 64; Indels 0; Gaps 0;

QY 182 GCTACGAATGTGTGGAGAGCCCACTGTTTTCGCGCTGATTTTTCGCGGCTCCACGAAA 241
DB 70 GCTACGTGGGGAGAGAGCCGCACTGTTTTCGCGGCTGATTTTTCGCGGCTCCACGAAA 129
QY 242 TCGTGTATCGCCCTCGTGTGTGGCGAGCGCCCTCGGCGCCATCGCGCTACACCAT 301
DB 130 CCTGTGTGGCGCGCTGTGTGTGGCGAGCGCCCTCGGCGCCATCGCGCTACACCAT 189
QY 302 TCGTGTATCGCCCTCGTGTGTGGCGAGCGCCCTCGGCGCCATCGCGCTACACCAT 316
DB 190 CGT 204


```
/ Sequence 983, Application US/10100683
/ GENERAL INFORMATION:
/ APPLICANT: Rosen, et al.
/ TITLE OF INVENTION: Human Secreted Proteins
/ FILE REFERENCE: PS900
/ CURRENT APPLICATION NUMBER: US/10/100,683
/ PRIOR FILING DATE: 2002-03-19
/ PRIOR APPLICATION NUMBER: US 60/040,162
/ PRIOR FILING DATE: 1997-03-07
/ PRIOR APPLICATION NUMBER: US 60/043,576
/ PRIOR FILING DATE: 1997-04-11
/ PRIOR APPLICATION NUMBER: US 60/047,601
/ PRIOR FILING DATE: 1997-05-23
/ PRIOR APPLICATION NUMBER: US 60/056,845
/ PRIOR FILING DATE: 1997-08-22
/ PRIOR APPLICATION NUMBER: US 60/043,580
/ PRIOR FILING DATE: 1997-04-11
/ PRIOR APPLICATION NUMBER: US 60/047,599
/ PRIOR FILING DATE: 1997-05-23
/ PRIOR APPLICATION NUMBER: US 60/056,664
/ PRIOR FILING DATE: 1997-08-22
/ PRIOR APPLICATION NUMBER: US 60/043,314
/ PRIOR FILING DATE: 1997-04-11
/ PRIOR APPLICATION NUMBER: US 60/047,632
/ PRIOR FILING DATE: 1997-05-23
/ PRIOR APPLICATION NUMBER: US 60/056,892
/ PRIOR FILING DATE: 1997-08-22
/ Remaining Prior Application data removed - See file wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 13468
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 983
/ LENGTH: 3173
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ US-10-100-683-983

Query Match          4.6%; Score 34.6; DB 6; Length 3173;
Best Local Similarity 75.4%; Pred. No. 0.79;
Matches 43; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 244 CTGCTATGCGCCTCTGTTGGGCGAGGCGCCCTGGGGCGCATGCGCTCACCA 300
DB 2360 CAGCTCCCGCGCCCGAGTGGGCTCAGGCGCTCTGGGCGCATGCGCTCACCA 2304

RESULT 6
US-10-100-683-11511
/ Sequence 11511, Application US/10100683
/ GENERAL INFORMATION:
/ APPLICANT: Rosen, et al.
/ TITLE OF INVENTION: Human Secreted Proteins
/ FILE REFERENCE: PS900
/ CURRENT APPLICATION NUMBER: US/10/100,683
/ PRIOR FILING DATE: 2002-03-19
/ PRIOR APPLICATION NUMBER: US 60/040,162
/ PRIOR FILING DATE: 1997-03-07
/ PRIOR APPLICATION NUMBER: US 60/043,576
/ PRIOR FILING DATE: 1997-04-11
/ PRIOR APPLICATION NUMBER: US 60/047,601
/ PRIOR FILING DATE: 1997-05-23
/ PRIOR APPLICATION NUMBER: US 60/056,845
/ PRIOR FILING DATE: 1997-08-22
/ PRIOR APPLICATION NUMBER: US 60/043,580
/ PRIOR FILING DATE: 1997-04-11
/ PRIOR APPLICATION NUMBER: US 60/047,599
/ PRIOR FILING DATE: 1997-05-23
/ PRIOR APPLICATION NUMBER: US 60/056,664
/ PRIOR FILING DATE: 1997-08-22
/ PRIOR APPLICATION NUMBER: US 60/043,314
/ PRIOR FILING DATE: 1997-04-11
/ PRIOR APPLICATION NUMBER: US 60/047,632
/ PRIOR FILING DATE: 1997-05-23
/ PRIOR APPLICATION NUMBER: US 60/056,892
```

```
/ PRIOR FILING DATE: 1997-08-22
/ Remaining Prior Application data removed - See file wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 13468
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 11511
/ LENGTH: 7071
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ US-10-100-683-11511

Query Match          4.6%; Score 34.6; DB 6; Length 7071;
Best Local Similarity 75.4%; Pred. No. 1.2;
Matches 43; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 244 CTGCTATGCGCCTCTGTTGGGCGAGGCGCCCTGGGGCGCATGCGCTCACCA 300
DB 785 CAGCTCCCGCGCCCGAGTGGGCTCAGGCGCTCTGGGCGCATGCGCTCACCA 841

RESULT 7
US-10-767-701-7545
/ Sequence 7545, Application US/10767701
/ GENERAL INFORMATION:
/ APPLICANT: Kovalic, David K.
/ APPLICANT: Zhou, Yihua
/ TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
/ FILE REFERENCE: 38-21(5335)B
/ CURRENT APPLICATION NUMBER: US/10/767,701
/ PRIOR FILING DATE: 2004-01-29
/ NUMBER OF SEQ ID NOS: 63128
/ SEQ ID NO 7545
/ LENGTH: 666
/ TYPE: DNA
/ ORGANISM: Sorghum bicolor
/ FEATURE:
/ OTHER INFORMATION: Clone ID: SORBI-28MAY03-CLUS92927_1
/ US-10-767-701-7545

Query Match          4.6%; Score 34.4; DB 6; Length 666;
Best Local Similarity 57.4%; Pred. No. 0.42;
Matches 62; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

QY 215 GCCTGATTTTGGCGGCTCCACCGAATGCTGCTGATCGCCCTGTGTGGGCGCAGCG 274
DB 31 GCCGACCAACACGAGGAGGAGCAGCCCGCTCATCAACCGTTGTGAGCGGCGCGCG 90

QY 275 CCTGGGGCGCATGCGCTCACCATTTGCTGTGAACCTTCGCCAGC 322
DB 91 CGCTGCGCTCCCTGCGTGGCGAGCGCGGTGCTGTGTGGTGGCGAG 138

RESULT 8
US-10-767-701-744/c
/ Sequence 744, Application US/10767701
/ GENERAL INFORMATION:
/ APPLICANT: Kovalic, David K.
/ APPLICANT: Zhou, Yihua
/ APPLICANT: Cao, Yongwei
/ TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
/ FILE REFERENCE: 38-21(5335)B
/ CURRENT APPLICATION NUMBER: US/10/767,701
/ PRIOR FILING DATE: 2004-01-29
/ NUMBER OF SEQ ID NOS: 63128
/ SEQ ID NO 744
/ LENGTH: 561
/ TYPE: DNA
/ ORGANISM: Sorghum bicolor
/ FEATURE:
/ OTHER INFORMATION: Clone ID: SORBI-28MAY03-CLUS124346_1
/ US-10-767-701-744
```

NUMBER OF SEQUENCES: 7310
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: GENOME THERAPEUTICS CORPORATION
 STREET: 100 Beaver Street
 CITY: Waltham
 STATE: Massachusetts
 COUNTRY: USA
 ZIP: 02354
 COMPUTER READABLE FORM:
 MEDIUM TYPE: CD-ROM ISO9660
 COMPUTER: PC
 OPERATING SYSTEM: <Unknown>
 SOFTWARE: ASCII
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/10/417,884A
 FILING DATE: 17-Apr-2003
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/09/107,532
 FILING DATE: 30-Jun-1998
 APPLICATION NUMBER: 60/085,598
 FILING DATE: 14 May 1998
 APPLICATION NUMBER: 60/051571
 FILING DATE: July 2, 1997
 ATTORNEY/AGENT INFORMATION:
 NAME: Ariniello, Pamela Deneke
 REGISTRATION NUMBER: 40,489
 REFERENCE/DOCKET NUMBER: GTC-012
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (781)893-5007
 TELEFAX: (781)893-8277
 INFORMATION FOR SEQ ID NO: 1266:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 768 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: circular
 MOLECULE TYPE: DNA (genomic)
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 ORIGINAL SOURCE:
 ORGANISM: Enterococcus faecium
 FEATURE:
 NAME/KEY: m1sc feature
 LOCATION: (B) LOCATION 1...768
 SEQUENCE DESCRIPTION: SEQ ID NO: 1266:
 US-10-417-884A-1266

Query Match
 Best Local Similarity 5.3%; Score 39.8; DB 6; Length 768;
 Matches 104; Conservative 0; Mismatches 107; Indels 0; Gaps 0;

QY 107 TAAAAACCTCCCTGCTGTCAGGTTGGCATGATGCCGATTGGTGTGCTCT 166
 DB 92 TCAAGACACCTTACCTACCGTTTCGTTATATGATGATGACATTGGCATTTGATTCG 151
 QY 167 TGGTATTCATATAGGCTACGATGATGAGGAGCCCACTGTTTCCGCGCTGATTTTCG 226
 DB 152 TTGGGAAAGCTGCCGATTTTCATCCACTAGTCGTCAGTTGATATCCCTCTGGTCTATG 211
 QY 227 CGGGCTCCAGGAATGCTGTGTCATCGCCCTGCTGTTGGGCGCAGCGCCCTGGGCGCA 286
 DB 212 CTGGTTGCGCCAAATTATCAAGTCAGTCATGCTTGCTAGTCACAGCCCAATTGTTTCCA 271
 QY 287 TCGGCGTCAACAATGCTGTGTAAGTCCG 317
 DB 272 TCGTTTCTCAACCTTTCTAGTCAATTCCG 302

RESULT 3
 US-10-767-701-5263
 SEQUENCE INFORMATION:
 GENERAL INFORMATION:
 APPLICANT: Kovalic, David K.

APPLICANT: Zhou, Yihua
 APPLICANT: Cao, Yongwei
 TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
 TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
 FILE REFERENCE: 38-21(53535)B
 CURRENT APPLICATION NUMBER: US/10/767,701
 CURRENT FILING DATE: 2004-01-29
 NUMBER OF SEQ ID NOS: 63128
 SEQ ID NO 5263
 LENGTH: 619
 TYPE: DNA
 ORGANISM: Sorghum bicolor
 FEATURE:
 OTHER INFORMATION: clone ID: SORBI-28MAY03-CLUS11974_1
 US-10-767-701-5263

Query Match
 Best Local Similarity 4.7%; Score 35.4; DB 6; Length 619;
 Matches 72; Conservative 0; Mismatches 61; Indels 0; Gaps 0;

QY 226 GCGGCTCCACCGAATGCTGTGTCATCGCCCTGCTGTGTCGCGCAGCGCCCTGGGCGC 285
 DB 132 GCGGCTTCACCTTGTGTGTCGTGTGTCGCGGCTGCGGCAAGTGTGCGCTC 191
 QY 286 ATCGGCTCACCATATGCTGTGTAATTCGCGCAGTATTCATGCTTTGATCCG 345
 DB 192 ATCGGCGAGCGCGTGTGTGATCTTCAAGCCCGTGTGTGCGGCGCATCAG 251
 QY 346 CTGCATGTGTCA 358
 DB 252 GTGGCGCTGCTCA 264

RESULT 4
 US-10-767-701-10866
 SEQUENCE INFORMATION:
 GENERAL INFORMATION:
 APPLICANT: Kovalic, David K.
 APPLICANT: Zhou, Yihua
 APPLICANT: Cao, Yongwei
 TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
 TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
 FILE REFERENCE: 38-21(53535)B
 CURRENT APPLICATION NUMBER: US/10/767,701
 CURRENT FILING DATE: 2004-01-29
 NUMBER OF SEQ ID NOS: 63128
 SEQ ID NO 10866
 LENGTH: 1034
 TYPE: DNA
 ORGANISM: Sorghum bicolor
 FEATURE:
 OTHER INFORMATION: clone ID: SORBI-28MAY03-CLUS4850_1
 US-10-767-701-10866

Query Match
 Best Local Similarity 4.6%; Score 34.6; DB 6; Length 1034;
 Matches 88; Conservative 0; Mismatches 89; Indels 0; Gaps 0;

QY 241 ATGCTGTGATGAGCTGTGTCGCGCAGCGCCCTGCGCGCCATGCGCTACACA 300
 DB 151 ATGGGCGGCCCTACTACAGCGGAGCCCTGCGCGCCAGCGCGCGGCCCT 210
 QY 301 TTGCTGTGATCTTCCGCACTATTCATGCTTTTATTCCTGCTGATGTGTA 360
 DB 211 GTGGCGGTGTGTCGCGGAGTTCGCGCGCGTGTGCTGCTTACGCTGACAG 270
 QY 361 AACCCATGCGCGTTTCTATTCGCTTTGCGCTTATGACAGAGCTTACAGCAGTC 417
 DB 271 AAGGCATAGCTCTCCAGCGGCACTTCAACCGTCACCGAGCCAGGCGCGC 327

RESULT 5
 US-10-100-683-983/c

APPLICANT: Day, Craig
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
FILE REFERENCE: 210121.427C9
CURRENT APPLICATION NUMBER: US/09/439,313
CURRENT FILING DATE: 1999-11-12
NUMBER OF SEQ ID NOS: 575
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO: 12
LENGTH: 751
TYPE: DNA
ORGANISM: Homo sapien
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(751)
OTHER INFORMATION: n = A,T,C or G
US-09-439-313-12

Query Match 4.8%; Score 36.4; DB 4; Length 751;
Best Local Similarity 59.8%; Pred. No. 0.11;
Matches 58; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

QY 227 CGGGCTCCACCGAATGCTGTCATGCGCCCTGTTGAGGCGAGCCCTGGGGCGCA 286
DB 410 CGTTCTTCTTCATCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCT 351
QY 287 TCGCGCTCACCAATGCTGCTGTAAGTTCCGCCAGCT 323
DB 350 TGGGTACACCAATGCTGAGCACTTCCCTGAGCT 314

RESULT 15

US-09-352-616A-12/C
Sequence 12, Application US/09352616A
Patent No. 6395278
GENERAL INFORMATION:
APPLICANT: Dillon, Davin C.
APPLICANT: Harlocker, Susan Louise
APPLICANT: Liang, Yugu
APPLICANT: Xu, Jiangchun
APPLICANT: Mitcham, Jennifer Lynn
TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS
FILE REFERENCE: 210121.427C8
CURRENT APPLICATION NUMBER: US/09/352,616A
CURRENT FILING DATE: 1999-07-13
NUMBER OF SEQ ID NOS: 472
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO: 12
LENGTH: 751
TYPE: DNA
ORGANISM: Homo sapien
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(751)
OTHER INFORMATION: n = A,T,C or G
US-09-352-616A-12

Query Match 4.8%; Score 36.4; DB 4; Length 751;
Best Local Similarity 59.8%; Pred. No. 0.11;
Matches 58; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

QY 227 CGGGCTCCACCGAATGCTGTCATGCGCCCTGTTGAGGCGAGCCCTGGGGCGCA 286
DB 410 CGTTCTTCTTCATCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCT 351
QY 287 TCGCGCTCACCAATGCTGCTGTAAGTTCCGCCAGCT 323
DB 350 TGGGTACACCAATGCTGAGCACTTCCCTGAGCT 314

Tue Apr 20 06:47:24 2004

us-10-608-504-6_copy_101_853.rn1

Page 6

PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 2761
LENGTH: 2061
TYPE: DNA
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-2761

Query Match 4.8%; Score 36.8; DB 4; Length 2061;
Best Local Similarity 59.8%; Pred. No. 0.11; 47; Indels 0; Gaps 0;
Matches 65; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

QY 211 TCCGGCTGATTTTCGGGGGCTCCACCGAATGCTGATCGCCCTCGTTG9GGCGCA 270
DB 1273 TACTTCTGATCACTCGCGCGCGCTCGCGCGCTGATCGGCTGACTCTCGCGCTC 1332
QY 271 GCGCCCTGCGCGCCATCGCGCTCACACATTCCTGCTGACTTCCGCCACG 322
DB 1333 GCGCCCGACGCGCGACCTTCTGATCGCGCGCTGCTACTTCTGCTACG 1384

RESULT 12
US-09-020-956-12/c
Sequence 12, Application US/09020956
Patent No. 6261562
GENERAL INFORMATION:
APPLICANT: Xu, Jiangchun
APPLICANT: Dillon, David C.
TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY OF PROSTATE CANCER AND METHODS FO
NUMBER OF SEQUENCES: 178
CORRESPONDENCE ADDRESS:
ADDRESS: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/020,956
FILING DATE: 09-FEB-1998
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Makl, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 210121.427C2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 751 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-09-020-956-12

Query Match 4.8%; Score 36.4; DB 3; Length 751;
Best Local Similarity 59.8%; Pred. No. 0.11; 39; Indels 0; Gaps 0;
Matches 58; Conservative 0; Mismatches 39; Indels 0; Gaps 0;
QY 227 CCGGCTCACCGAATGCTGATCGCCCTGTTG9GGCGCGCGCCCTGGGGCGCA 286
DB 410 CGTTCTTCTCATCCTCTCTCATCTTCATCTGAGTNGCAGCTGCTGCTGCT 351
QY 287 TCGCGCTCACCATTTGCTGTGTAATTCGCGCAGCT 323

DB 350 TGGTACACCAATGCTGAGACCTTCCCTGACGT 314

RESULT 13
US-09-030-607-12/c
Sequence 12, Application US/09030607
Patent No. 6262245
GENERAL INFORMATION:
APPLICANT: Xu, Jiangchun
APPLICANT: Dillon, David C.
TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY OF PROSTATE CANCER AND METHODS FO
NUMBER OF SEQUENCES: 224
CORRESPONDENCE ADDRESS:
ADDRESS: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/030,607
FILING DATE: 25-FEB-1998
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Makl, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 210121.427C3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 751 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-09-030-607-12

Query Match 4.8%; Score 36.4; DB 3; Length 751;
Best Local Similarity 59.8%; Pred. No. 0.11; 39; Indels 0; Gaps 0;
Matches 58; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

QY 227 CCGGCTCACCGAATGCTGATCGCCCTGTTG9GGCGCGCGCCCTGGGGCGCA 286
DB 410 CGTTCTTCTCATCCTCTCTCATCTTCATCTGAGTNGCAGCTGCTGCTGCT 351
QY 287 TCGCGCTCACCATTTGCTGTGTAATTCGCGCAGCT 323
DB 350 TGGTACACCAATGCTGAGACCTTCCCTGACGT 314

RESULT 14
US-09-439-313-12/c
Sequence 12, Application US/09439313
Patent No. 6329505
GENERAL INFORMATION:
APPLICANT: Xu, Jiangchun
APPLICANT: Dillon, David C.
APPLICANT: Mitcham, Jennifer L.
APPLICANT: Harlocker, Susan Louise
APPLICANT: Jiang Yuqi
APPLICANT: Reed, Steven G.
APPLICANT: Kales, Michael
APPLICANT: Ranger, Gary
APPLICANT: Retter, Mark
APPLICANT: Solik, John

US-09-252-991A-11376

Query Match 5.2%; Score 39.2; DB 4; Length 1038;
Best Local Similarity 55.0%; Pred. No. 0.017;
Matches 77; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

QY 209 TTTCGGGCTGATTTTCGGGGCTCCACGAAATGCTGATTCGCTCGTTGGGGCG 268
DB 473 TGTCCGCCATCGCTTCCTCGCGCGCGCCGCAAGTGTGCGCATGCTCAAGGGCG 532
QY 269 CAGCGCGCTGGGGCGGCATCGGCTCACCACATGCTGGTGAAGTCCGCCACGTAATCT 328
DB 533 GTGCGGGGTTCTTTTGATCATCTTACACACCTCTCTCTGACTCGACGACCTGCTCT 592
QY 329 ATGCGTTTCATTCGCGCTG 348
DB 593 ACGGCATGAGCTTGCGGCGG 612

RESULT 8

US-09-252-991A-11418

Sequence 11418, Application US/09252991A

Patent No. 6551795

GENERAL INFORMATION:

APPLICANT: Marc J. Rubenfield et al.

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

FILE REFERENCE: 107196.136

CURRENT APPLICATION NUMBER: US/09/252,991A

PRIOR FILING DATE: 1999-02-18

PRIOR APPLICATION NUMBER: US 60/074,788

PRIOR FILING DATE: 1998-02-18

PRIOR APPLICATION NUMBER: US 60/094,190

PRIOR FILING DATE: 1998-07-27

NUMBER OF SEQ ID NOS: 33142

SEQ ID NO 11418

LENGTH: 1086

TYPE: DNA

ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-11418

Query Match 5.2%; Score 39.2; DB 4; Length 1086;
Best Local Similarity 55.0%; Pred. No. 0.017;
Matches 77; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

QY 209 TTTCGGGCTGATTTTCGGGGCTCCACGAAATGCTGATTCGCTCGTTGGGGCG 268
DB 222 TGTCCGCCATCGCTTCCTCGCGCGCGCCGCAAGTGTGCGCATGCTCAAGGGCG 281
QY 269 CAGCGCGCTGGGGCGGCATCGGCTCACCACATGCTGGTGAAGTCCGCCACGTAATCT 328
DB 282 GTGCGGGGTTCTTTTGATCATCTTACACACCTCTCTCTGACTCGACGACCTGCTCT 341
QY 329 ATGCGTTTCATTCGCGCTG 348
DB 342 ACGGCATGAGCTTGCGGCGG 361

RESULT 9

US-09-621-976-8976/c

Sequence 8976, Application US/09621976

Patent No. 6639063

GENERAL INFORMATION:

APPLICANT: Dumas Milne Edwards, J. B.

APPLICANT: Jobert, S.

TITLE OF INVENTION: ESTS and Encoded Human Proteins.

FILE REFERENCE: GENSET.054PR2

CURRENT APPLICATION NUMBER: US/09/621,976

CURRENT FILING DATE: 2000-07-21

NUMBER OF SEQ ID NOS: 19335

SOFTWARE: Patent.pm

SEQ ID NO 8976

LENGTH: 399

TYPE: DNA
ORGANISM: Homo sapiens
US-09-621-976-8976

Query Match 5.2%; Score 39; DB 4; Length 399;
Best Local Similarity 12.6%; Pred. No. 0.012;
Matches 30; Conservative 112; Mismatches 97; Indels 0; Gaps 0;

QY 509 GAGTGGCGATCGAGAGTTGATTCCTTTGAAATTAAGGCGCTCGAGTCCCTTGTCT 568
DB 242 SAMWRKKKMTCKMGRSWSGRSTGYAAMWKSMCTSRKMYKKRKRKRSTKRT 183
QY 569 CTCTCTTGTACAGCTGACTTGGATTCCCGCGAAGAAAGCAGATCCCTCTGCG 628
DB 182 CYRSTYKWKAYTKRRKRWTKYTYKSYMSKKTWEMKTAIYWKMTKRTKW 123
QY 629 TGCTCGAGGTTGAGCTTCACCATTCCTCTTGTAATTCAGGTCAGGCGCTATTG 668
DB 122 CTGCKCTTYMAGTMMYRBYRYAKRAKSKRCITWSTCYCMKMAKCSYWSMS 63
QY 689 CGGCGCTGATCTTTGGGCTGTGACCATCGGTACTCTCTTGGGAAAGGCT 747
DB 62 MKWKSMWKMYTYYYKMSKMTYMSKCYAKRCRTAKTYMTWTGRT 4

RESULT 10

US-09-252-991A-3205/c

Sequence 3205, Application US/09252991A

Patent No. 6551795

GENERAL INFORMATION:

APPLICANT: Marc J. Rubenfield et al.

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

FILE REFERENCE: 107196.136

CURRENT APPLICATION NUMBER: US/09/252,991A

PRIOR FILING DATE: 1999-02-18

PRIOR APPLICATION NUMBER: US 60/074,788

PRIOR FILING DATE: 1998-02-18

PRIOR APPLICATION NUMBER: US 60/094,190

PRIOR FILING DATE: 1998-07-27

NUMBER OF SEQ ID NOS: 33142

SEQ ID NO 3205

LENGTH: 1140

TYPE: DNA

ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-3205

Query Match 4.9%; Score 36.8; DB 4; Length 1140;
Best Local Similarity 58.0%; Pred. No. 0.1;
Matches 65; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

QY 211 TCCGCGCTGATTTTCGGGGCTCCACGAAATGCTGATTCGCTCGTTGGGGCGCA 270
DB 465 TACTTCGTGATCAACCTGCGCGCGCCCTCGCGCGCTGATCGGCTGACTCGGGCTC 406
QY 271 GCGCGCTGGGGCGGCATCGGCTCACCACATGCTGGTGAAGTCCGCCGCG 322
DB 405 GCGCGCGAGCGCGGACCTTCTGATCAGCGCGCTGATCTGCTGACG 354

RESULT 11

US-09-252-991A-2761

Sequence 2761, Application US/09252991A

Patent No. 6551795

GENERAL INFORMATION:

APPLICANT: Marc J. Rubenfield et al.

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

FILE REFERENCE: 107196.136

CURRENT APPLICATION NUMBER: US/09/252,991A

CURRENT FILING DATE: 1999-02-18

PRIOR APPLICATION NUMBER: US 60/074,788

FILING DATE: 26-Aug-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 IMMU
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-9300
TELEFAX: (703) 683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: pTZgpt-Fls
US-08-232-463-14

Query Match 6.1%; Score 46.2; DB 1; Length 7218;
Best Local Similarity 5.2%; Pred. No. 0.00027;
Matches 21; Conservative 214; Mismatches 172; Indels 0; Gaps 0;

QY 215 GCGTATTTGCGCGCTCCACGAAATGCTGATCGCCCTGTTGCGCGCAGCG 274
DB 1041 GCGTATTTGCGCGCTCCACGAAATGCTGATCGCCCTGTTGCGCGCAGCG 1100
QY 275 CCTGGGCGCATGCGCTCACCACATTCCTGCTGAACTCCGCCACGATTCATGCT 334
DB 1101 YY 1160
QY 335 TTTCATTCCTGCTGATGCTGCTCAAAACCCATTCGCTTCTATGCTTTCGCGC 394
DB 1161 YY 1220
QY 395 TTATCGAGAGCTGACGATCTGCGGCGAGCCCGCAGGCTGCTGCGCTGCGCAGC 454
DB 1221 YY 1280
QY 455 TTATCGATGCAATAGCGTTTCATCTCTACTGCGGATTCGCGGCTCACCAGAGTGG 514
DB 1281 YY 1340
QY 515 CGATCGAGAGTATTCCTTTGAATTAAGGCGCTCGAGTTCGCCCTTGTCTCTCT 574
DB 1341 YY 1400
QY 575 TTGTCAGCGCTGATTCGATTCCTGCGCAAGAGAGATTCCT 621
DB 1401 YY 1447

RESULT 6
US-09-107-532A-1266
Sequence 1266, Application US/09107532A
Patent No. 6583275

GENERAL INFORMATION:
APPLICANT: Lyth A Doucette-Stamm and David Bush
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
NUMBER OF SEQUENCES: 7210
CORRESPONDENCE ADDRESS:
ADDRESS: GENOME THERAPEUTICS CORPORATION
STREET: 100 Beaver Street
CITY: Waltham
STATE: Massachusetts
COUNTRY: USA
ZIP: 02354
COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660
COMPUTER: PC
OPERATING SYSTEM: <Unknown>
SOFTWARE: ASCII

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/107,532A
FILING DATE: 30-Jun-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/085,598
FILING DATE: 14 May 1998
APPLICATION NUMBER: 60/051571
FILING DATE: July 2, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Arinello, Pamela Deneke
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-012
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781) 893-5007
TELEFAX: (781) 893-8277
INFORMATION FOR SEQ ID NO: 1266:
SEQUENCE CHARACTERISTICS:
LENGTH: 768 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Enterococcus faecium
FEATURE:
NAME/KEY: misc feature
LOCATION: (B) LOCATION 1...768
SEQUENCE DESCRIPTION: SEQ ID NO: 1266:
US-09-107-532A-1266

Query Match 5.3%; Score 39.8; DB 4; Length 768;
Best Local Similarity 4.93%; Pred. No. 0.0092;
Matches 104; Conservative 0; Mismatches 107; Indels 0; Gaps 0;

QY 107 TAAACCTCCCTGCTGCGAGTTTGCGCATGACCCGATTTGCGCTGATTCCT 166
DB 92 TCAAGACAGCTTACCTACCGCTTTCGTTATGATGATTCGATTCGATTCGATTCG 151
QY 167 TGGTATTCATATGCGCTGCGAATGCTGCGAGCCCGCATGTTTCCGCGCTGATTCG 226
DB 152 TTGGGAAGCTGCGGATTTTCATCCATGATGCTACGTTAGTCCGCTGCTGATG 211
QY 227 CGGCTCACCAGAAATGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 286
DB 212 CTGCTTCGCGCAATTCATACAGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCT 271
QY 287 TCGGCTCACCAGATTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 317
DB 272 TCGTTTCTGCGACCTTCTAGTCAATTCGCG 302

RESULT 7
US-09-252-991A-11376
Sequence 11376, Application US/09252991A
Patent No. 6551795

GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 11376
LENGTH: 1038
TYPE: DNA
ORGANISM: Pseudomonas aeruginosa

SEQ ID NO 6
LENGTH: 1271
TYPE: DNA
ORGANISM: Corynebacterium glutamicum
FEATURE:
NAME/KEY: gene
LOCATION: (101)..(853)
OTHER INFORMATION: bmf
NAME/KEY: gene
LOCATION: (853)..(1176)
OTHER INFORMATION: bme
OTHER INFORMATION: ATCC13032
US-09-471-803A-6

Query Match 100.0%; Score 753; DB 4; Length 1271;
Best Local Similarity 100.0%; Pred. No. 2.9e-211;
Matches 753; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTGCAAAAACGCAAGATTCTTCAAGCTGAGGTGTGCCATTCAGGCGGCTG 60
DB 101 GTGCAAAAACGCAAGATTCTTCAAGCTGAGGTGTGCCATTCAGGCGGCTG 160
QY 61 GAAACGATGATAAAGTTATGAGGCTACGAAATCGGCAAGGTCTAATAACCTCCCTT 120
DB 161 GAAACGATGATAAAGTTATGAGGCTACGAAATCGGCAAGGTCTAATAACCTCCCTT 220
QY 121 GCTGAGGTTTGGGCAATGATCCGATGTGATGTGCTTGTCTTGTATTCATATC 180
DB 221 GCTGAGGTTTGGGCAATGATCCGATGTGATGTGCTTGTCTTGTATTCATATC 280
QY 181 GGTGATGATGTGGGCAAGGCTGTTTCCGCGCTGATTTTGGGCGGCTCCACCGAA 240
DB 281 GGTGATGATGTGGGCAAGGCTGTTTCCGCGCTGATTTTGGGCGGCTCCACCGAA 340
QY 241 ATGCTGATGATGCGCTCGTTGTGGGCGGAGCGCCCTGAGGCGCATCGCGTCAACCA 300
DB 341 ATGCTGATGATGCGCTCGTTGTGGGCGGAGCGCCCTGAGGCGCATCGCGTCAACCA 400
QY 301 TTGCTGTGATCTTCGCGCAAGTATTCATGCTTTTCAATCCGCTGATGTGTCAA 360
DB 401 TTGCTGTGATCTTCGCGCAAGTATTCATGCTTTTCAATCCGCTGATGTGTCAA 460
QY 361 AACCCATGCGCGCTTCTTATTCGCTTTCGCGCTTATGAGGAGCTACGAGTCACT 420
DB 461 AACCCATGCGCGCTTCTTATTCGCTTTCGCGCTTATGAGGAGCTACGAGTCACT 520
QY 421 GCGGCGAGGCGCGAGGCTGTGTGGGCGGCACTTATTCATGCAATAGCGTTTCA 480
DB 521 GCGGCGAGGCGCGAGGCTGTGTGGGCGGCACTTATTCATGCAATAGCGTTTCA 580
QY 481 TCCCTATGCTGATTCGCGCGCTTCACCGAGGTGGGATGCGAGGTGATTCCTTTGA 540
DB 581 TCCCTATGCTGATTCGCGCGCTTCACCGAGGTGGGATGCGAGGTGATTCCTTTGA 640
QY 541 ATTAAGGCGCTCGAGTTCGCGCTTGTCTCTCTTGTGACGCTGACTTGGATTCTGC 600
DB 641 ATTAAGGCGCTCGAGTTCGCGCTTGTCTCTCTTGTGACGCTGACTTGGATTCTGC 700
QY 601 CGAAGCAAAAAGCATTCCTTCTCTGCTGCTGCGAGTTTGAAGTTCAACATGCTCT 660
DB 701 CGAAGCAAAAAGCATTCCTTCTCTGCTGCTGCGAGTTTGAAGTTCAACATGCTCT 760
QY 661 GTGATATTCAGAGTCAAGGCGCTATTTGCGCGCTGCTGATCTTCTGCGCTGTGACC 720
DB 761 GTGATATTCAGAGTCAAGGCGCTATTTGCGCGCTGCTGATCTTCTGCGCTGTGACC 820
QY 721 ATCCGATCTCTTCTTGGGAAAGGCTGCTAAA 753
DB 821 ATCCGATCTCTTCTTGGGAAAGGCTGCTAAA 853

RESULT 4
US-09-489-039A-3016

Sequence 3016, Application US/09489039A
Patent No. 6610836
GENERAL INFORMATION:
APPLICANT: Gary Breton et. al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 2709.2004001
CURRENT APPLICATION NUMBER: US/09/489, 039A
CURRENT FILING DATE: 2000-01-27
PRIOR APPLICATION NUMBER: US 60/117,747
PRIOR FILING DATE: 1999-01-29
NUMBER OF SEQ ID NOS: 14342
SEQ ID NO 3016
LENGTH: 756
TYPE: DNA
ORGANISM: Klebsiella pneumoniae
US-09-489-039A-3016

Query Match 7.8%; Score 58.6; DB 4; Length 756;
Best Local Similarity 52.2%; Pred. No. 8.6e-09;
Matches 130; Conservative 0; Mismatches 119; Indels 0; Gaps 0;

QY 95 TCGCGAAGGCTTAAACCTCCCTGCTGACAGTTTGGGCAATGACCCGATTTGATTG 154
DB 62 TCGCGAAGGCTTAAACCTCCCTGCTGACAGTTTGGGCAATGACCCGATTTGATTG 121
QY 155 GCTTGTCTCTTGTATTCATATGAGGCTTACGAAATGTGGGAGGCCCACTGTTTCCG 214
DB 122 CTTGGAGCTCAATGCTACCGGCTGAGCTTCAACCCCGCTGAAAGCTGTTTCTTCT 181
QY 215 GCTGATTTTGGGCGGCTCCACCGAAATGCTGTGATGCGCTTGTGTGGGCGAGCG 274
DB 182 GCATTTATTCAGCGCGGCGGCGGCAAGTGTGTATCACCGCATGTTGGCGCGGAGTT 241
QY 275 CCTGCGCGCATGCGGCTCAACCATGCTGTGTGAACTTCGCGCATTTTATGCGT 334
DB 242 CGCTGTGGGTGCGCGCGCTTCAAGGATGATGCGATGATGCTCCGACGAGCTATGCG 301
QY 335 TTTCATTC 343
DB 302 CTTCATTC 310

RESULT 5
US-08-232-463-14
Sequence 14, Application US/08232463
Patent No. 5670367
GENERAL INFORMATION:
APPLICANT: DORNER, F.
APPLICANT: SCHRIFFLINER, F.
APPLICANT: FALKNER, F. G.
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,463
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935,313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6

QY 241 ATGCTGATCGCCCTGCTGTTGGGCGAGGCGCCCTGAGGCGCATCGCCCTACACCA 300
DB 241 ATGCTGATCGCCCTGCTGTTGGGCGAGGCGCCCTGAGGCGCATCGCCCTACACCA 300
QY 301 TTGCTGATGAACTTCGCGCAGATTCATATGCGTTTCAATCCGCTGATGTTGTTCAA 360
DB 301 TTGCTGATGAACTTCGCGCAGATTCATATGCGTTTCAATCCGCTGATGTTGTTCAA 360
QY 361 AACCCATGTCGCGCTTCTATTCGCTTTCGCGCTTTCGCGCTTTCGCGCTTTCGCGCT 420
DB 361 AACCCATGTCGCGCTTCTATTCGCTTTCGCGCTTTCGCGCTTTCGCGCTTTCGCGCT 420
QY 421 GCGGCGAGGCGCGCAGCTGCTGCTGCGGCGGCGATTCATTCGAAATAGCGTTTCA 480
DB 421 GCGGCGAGGCGCGCAGCTGCTGCTGCGGCGGCGATTCATTCGAAATAGCGTTTCA 480
QY 481 TCTCTACTGAGTATTCGCGCGCTCTCAACCGAGTGGCGATTCGAGTGAATTCCTTTGAA 540
DB 481 TCTCTACTGAGTATTCGCGCGCTCTCAACCGAGTGGCGATTCGAGTGAATTCCTTTGAA 540
QY 541 ATTAAGGCGCTGAGTTCGCGCTTTCGCTCTCTCTTTCGCGCTGATTCGCTTTCGCG 600
DB 541 ATTAAGGCGCTGAGTTCGCGCTTTCGCTCTCTCTTTCGCGCTGATTCGCTTTCGCG 600
QY 601 CGAAGCAAAAAGAGATCCCTTCTCTGCTGCTGCGAGTTTGAAGTTTCACTTCTCT 660
DB 601 CGAAGCAAAAAGAGATCCCTTCTCTGCTGCTGCGAGTTTGAAGTTTCACTTCTCT 660
QY 661 GTGGTAATTCAGATTCAGGCGCTTTCGCGGCGCTGATTCGCTTTCGCGGCTTTCG 720
DB 661 GTGGTAATTCAGATTCAGGCGCTTTCGCGGCGCTGATTCGCTTTCGCGGCTTTCG 720
QY 721 ATCCGTAATTCCTTCTTGGGAAAAGCTGCTAAA 753
DB 721 ATCCGTAATTCCTTCTTGGGAAAAGCTGCTAAA 753

RESULT 2
US-09-471-803A-1
/ Sequence 1, Application US/09471803A
/ Patent No. 6613545
/ GENERAL INFORMATION:
/ APPLICANT: KENNERKNECHT, NICOLE
/ APPLICANT: SAHM, HERMANN
/ APPLICANT: EGSELING, LOTMAR
/ APPLICANT: PFEFFERLE, WALTER
/ TITLE OF INVENTION: NUCLEOTIDE SEQUENCE CODING FOR THE EXPORT OF
/ TITLE OF INVENTION: BRANCHED CHAIN AMINO ACIDS, PROCESS FOR THE
/ FILE REFERENCE: 21123/265496/MAS
/ CURRENT APPLICATION NUMBER: US/09/471, 803A
/ PRIOR FILING DATE: 1999-12-23
/ PRIOR APPLICATION NUMBER: DE 199 51 708.8
/ NUMBER OF SEQ ID NOS: 12
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 1
/ LENGTH: 1271
/ TYPE: DNA
/ ORGANISM: Corynebacterium glutamicum
/ FEATURE:
/ NAME/KEY: gene
/ LOCATION: (101)..(853)
/ OTHER INFORMATION: brnf
/ NAME/KEY: gene
/ LOCATION: (853)..(1176)
/ OTHER INFORMATION: brne
/ OTHER INFORMATION: ATCC14752
US-09-471-803A-1
Query Match 100.0%; Score 753; DB 4; Length 1271;
Best Local Similarity 100.0%; Pred. No. 2,9e-231;
Matches 753; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTCAAAAAAAGCAGAGATTCATTCAGCTGAGAGTGTGCGCATTCAGAGGACGCTTG 60
DB 101 GTCAAAAAAAGCAGAGATTCATTCAGCTGAGAGTGTGCGCATTCAGAGGACGCTTG 160
QY 61 GAACCGAGATTAAGATTCAGGCTTCAGAAATGCGCAAGTCAAAACCTCCCTT 120
DB 161 GAACCGAGATTAAGATTCAGGCTTCAGAAATGCGCAAGTCAAAACCTCCCTT 220
QY 121 GCTGAGATTCAGGCTTCAGGCTTCAGGCTTCAGGCTTCAGGCTTCAGGCTTCAGGCT 180
DB 221 GCTGAGATTCAGGCTTCAGGCTTCAGGCTTCAGGCTTCAGGCTTCAGGCTTCAGGCT 280
QY 181 GGTACGATGATGAGGAGCGCCCACTGTTTTCGCGCTGATTTTCGCGGCTTCACCGGA 240
DB 281 GGTACGATGATGAGGAGCGCCCACTGTTTTCGCGCTGATTTTCGCGGCTTCACCGGA 340
QY 241 ATGCTGATTCAGGCTTCAGGCTTCAGGCTTCAGGCTTCAGGCTTCAGGCTTCAGGCT 300
DB 341 ATGCTGATTCAGGCTTCAGGCTTCAGGCTTCAGGCTTCAGGCTTCAGGCTTCAGGCT 400
QY 301 TTGCTGATGAACTTCGCGCAGATTCATATGCGTTTCAATCCGCTGATGTTGTTCAA 360
DB 401 TTGCTGATGAACTTCGCGCAGATTCATATGCGTTTCAATCCGCTGATGTTGTTCAA 460
QY 361 AACCCATTCGCGCTTCTATTCGCTTTCGCGCTTTCGCGCTTTCGCGCTTTCGCGCT 420
DB 461 AACCCATTCGCGCTTCTATTCGCTTTCGCGCTTTCGCGCTTTCGCGCTTTCGCGCT 520
QY 421 GCGGCGAGGCGCGCAGCTGCTGCTGCGGCGGCGATTCATTCGAAATAGCGTTTCA 480
DB 521 GCGGCGAGGCGCGCAGCTGCTGCTGCGGCGGCGATTCATTCGAAATAGCGTTTCA 580
QY 481 TCTCTACTGAGTATTCGCGCGCTCTCAACCGAGTGGCGATTCGAGTGAATTCCTTTGAA 540
DB 581 TCTCTACTGAGTATTCGCGCGCTCTCAACCGAGTGGCGATTCGAGTGAATTCCTTTGAA 640
QY 541 ATTAAGGCGCTGAGTTCGCGCTTTCGCTCTCTCTTTCGCGCTGATTCGCTTTCGCG 600
DB 641 ATTAAGGCGCTGAGTTCGCGCTTTCGCTCTCTCTTTCGCGCTGATTCGCTTTCGCG 700
QY 601 CGAAGCAAAAAGAGATCCCTTCTCTGCTGCTGCGAGTTTGAAGTTTCACTTCTCT 660
DB 701 CGAAGCAAAAAGAGATCCCTTCTCTGCTGCTGCGAGTTTGAAGTTTCACTTCTCT 760
QY 661 GTGGTAATTCAGATTCAGGCGCTTTCGCGGCGCTGATTCGCTTTCGCGGCTTTCG 720
DB 761 GTGGTAATTCAGATTCAGGCGCTTTCGCGGCGCTGATTCGCTTTCGCGGCTTTCG 820
QY 721 ATCCGTAATTCCTTCTTGGGAAAAGCTGCTAAA 753
DB 821 ATCCGTAATTCCTTCTTGGGAAAAGCTGCTAAA 853

RESULT 3
US-09-471-803A-6
/ Sequence 6, Application US/09471803A
/ Patent No. 6613545
/ GENERAL INFORMATION:
/ APPLICANT: KENNERKNECHT, NICOLE
/ APPLICANT: SAHM, HERMANN
/ APPLICANT: EGSELING, LOTMAR
/ APPLICANT: PFEFFERLE, WALTER
/ TITLE OF INVENTION: NUCLEOTIDE SEQUENCE CODING FOR THE EXPORT OF
/ TITLE OF INVENTION: BRANCHED CHAIN AMINO ACIDS, PROCESS FOR THE
/ FILE REFERENCE: 21123/265496/MAS
/ CURRENT APPLICATION NUMBER: US/09/471, 803A
/ PRIOR FILING DATE: 1999-12-23
/ PRIOR APPLICATION NUMBER: DE 199 51 708.8
/ NUMBER OF SEQ ID NOS: 12
/ SOFTWARE: PatentIn Ver. 2.1

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 19, 2004, 10:57:33 ; Search time 63.4023 Seconds

(without alignments)
6590.898 Million cell updates/sec

Title: US-10-608-504-6_COPY_101_853

Perfect score: 753

Sequence: 1 gtgcataaacgcaagagat.....tctcggaagagctgcctaa 753

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 682709 segs, 27747546 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: Issued Patents NA:*

1: /cgn2_6/prodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/prodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/prodata/2/ina/5A_COMB.seq:*
4: /cgn2_6/prodata/2/ina/5B_COMB.seq:*
5: /cgn2_6/prodata/2/ina/5A_COMB.seq:*
6: /cgn2_6/prodata/2/ina/5B_COMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	753	100.0	753	US-09-471-803A-2	Sequence 2, Appli
2	753	100.0	1271	US-09-471-803A-1	Sequence 1, Appli
3	753	100.0	1271	US-09-471-803A-6	Sequence 6, Appli
4	58.6	7.8	756	US-09-489-039A-3016	Sequence 3016, Ap
5	46.2	6.1	7218	US-08-232-463-14	Sequence 14, Appl
6	39.8	5.3	768	US-09-107-532A-1266	Sequence 1266, Ap
7	39.2	5.2	1038	US-09-252-991A-11376	Sequence 11376, A
8	39.2	5.2	1086	US-09-252-991A-11418	Sequence 11418, A
9	39.2	5.2	1086	US-09-252-991A-11418	Sequence 11418, A
10	36.8	4.9	1140	US-09-252-991A-3205	Sequence 3205, Ap
11	36.8	4.9	1140	US-09-252-991A-3205	Sequence 3205, Ap
12	36.4	4.8	751	US-09-020-956-12	Sequence 12, Appl
13	36.4	4.8	751	US-09-030-607-12	Sequence 12, Appl
14	36.4	4.8	751	US-09-439-313-12	Sequence 12, Appl
15	36.4	4.8	751	US-09-352-616A-12	Sequence 12, Appl
16	36.4	4.8	751	US-09-352-616A-12	Sequence 12, Appl
17	36.4	4.8	751	US-09-159-812-12	Sequence 12, Appl
18	36.4	4.8	751	US-09-636-215-12	Sequence 12, Appl
19	36.4	4.8	751	US-09-685-166A-12	Sequence 12, Appl
20	36.4	4.8	751	US-09-115-453-12	Sequence 12, Appl
21	36.4	4.8	751	US-09-688-489-12	Sequence 12, Appl
22	35.4	4.7	430	US-09-621-976-16656	Sequence 16656, A
23	35.4	4.7	474	US-09-621-976-18033	Sequence 18033, A
24	35.2	4.7	801	US-09-328-352-2399	Sequence 2399, Ap
25	35.2	4.7	4403765	US-09-103-840A-2	Sequence 1, Appli
26	35.2	4.7	441529	US-09-103-840A-1	Sequence 1, Appli
27	34.8	4.6	867	US-09-489-039A-5047	Sequence 5047, Ap

28	34.6	4.6	666	2	US-08-875-034A-1	Sequence 1, Appli
29	34.4	4.6	744	4	US-09-489-039A-1307	Sequence 1307, Ap
30	34.4	4.6	1224	4	US-09-266-965-22	Sequence 22, Appl
31	34.4	4.6	12249	4	US-09-266-965-74	Sequence 74, Appl
32	34.4	4.6	18331	4	US-09-266-965-96	Sequence 96, Appl
33	33.4	4.4	1875	3	US-09-422-869-21	Sequence 21, Appl
34	33.4	4.4	2949	3	US-08-433-522A-1	Sequence 1, Appli
35	33.4	4.4	2949	3	US-09-135-166-1	Sequence 1, Appli
36	33.4	4.4	2949	3	US-08-942-046-1	Sequence 1, Appli
37	33.4	4.4	2950	3	US-08-942-046-5	Sequence 5, Appli
38	33.4	4.4	2950	3	US-09-135-166-5	Sequence 5, Appli
39	33.4	4.4	2950	3	US-08-942-046-5	Sequence 5, Appli
40	33.4	4.4	2984	3	US-08-433-522A-3	Sequence 3, Appli
41	33.4	4.4	2984	3	US-09-135-166-3	Sequence 3, Appli
42	33.4	4.4	2984	3	US-08-942-046-3	Sequence 3, Appli
43	33.4	4.4	2987	3	US-08-433-522A-55	Sequence 55, Appli
44	33.4	4.4	2987	3	US-09-135-166-55	Sequence 55, Appli
45	33.4	4.4	2987	3	US-08-942-046-55	Sequence 55, Appli

ALIGNMENTS

RESULT 1						
US-09-471-803A-2						
Sequence 2, Application US/09471803A						
Parent No. 613545						
GENERAL INFORMATION:						
APPLICANT: KENNERNBECHT, NICOLE						
APPLICANT: SAMM, HERMANN						
APPLICANT: EGSELMING, LOTHAR						
TITLE OF INVENTION: NUCLEOTIDE SEQUENCE FOR THE EXPORT OF						
TITLE OF INVENTION: BRANCHED CHAIN AMINO ACIDS. PROCESS FOR THE						
FILE REFERENCE: 21123/265496/MAS						
CURRENT APPLICATION NUMBER: US/09/471,803A						
CURRENT FILING DATE: 1999-12-23						
PRIOR APPLICATION NUMBER: DE 199 51 708.8						
PRIOR FILING DATE: 1999-10-27						
NUMBER OF SEQ ID NOS: 12						
SOFTWARE: Patent Ver. 2.1						
SEQ ID NO 2						
LENGTH: 753						
TYPE: DNA						
ORGANISM: Corynebacterium glutamicum						
FEATURE:						
NAME/KEY: CDS						
LOCATION: (1)..(753)						
OTHER INFORMATION: DINF						
OTHER INFORMATION: ATCC14752						
US-09-471-803A-2						
Query Match						
Best Local Similarity 100.0%; Score 753; DB 4; Length 753;						
Matches 753; Conservative 0; Mismatches 0; Indels 0; Gaps 0;						
QY	1	GTGCAAAAACGCAAGATTTCATTAAGCCTGGAGGTGTGCCATTCAGGACGCTTG	60			
DB	1	GTGCAAAAACGCAAGATTTCATTAAGCCTGGAGGTGTGCCATTCAGGACGCTTG	60			
QY	61	GAACAGATGATTAAGTTATCGGCGGTACGAATTCGGGCAAGGCTTAATAAAGCTCCCTT	120			
DB	61	GAACAGATGATTAAGTTATCGGCGGTACGAATTCGGGCAAGGCTTAATAAAGCTCCCTT	120			
QY	121	GCTGAGGTTGGGATGATACCCGATTTGATTCGTTGCTCTTGGTTATTAATAC	180			
DB	121	GCTGAGGTTGGGATGATACCCGATTTGATTCGTTGCTCTTGGTTATTAATAC	180			
QY	181	GCTACGAATGATGAGGACGCGCCGCTTTTCGCGCTGATTTTCGCGGCTCCACCGAA	240			
DB	181	GCTACGAATGATGAGGACGCGCCGCTTTTCGCGCTGATTTTCGCGGCTCCACCGAA	240			

Db 18 YCWMSSGYSMWCTSGTY 2

RESULT 15

AAA10594/C
ID AAA10594 standard, DNA, 10732 BP.

AC AAA10594;

DT 29-JUN-2000 (first entry)

DE Gene encoding a subunit of cellulose synthase.

KW Cellulose synthase; cellulose production; increase yield; ds.

OS Vigna angularis.

PN JP2000060568-A.

PD 29-FEB-2000.

PF 26-AUG-1998; 98UP-00239998.

PR 26-AUG-1998; 98UP-00239998.

PA (MIZU) MIZUNO K.
(OUIP) OUI PAPER CO.

XX WPI; 2000-342371/30.

DR P-PSDB; AAY85179.

PT A gene encoding a cellulose synthetic equipment - for the improvement in
the amount of cellulose synthesized in a plant body.

PS Claim 2; Page 14-21; 32pp; Japanese.

XX This sequence represents a gene encoding a subunit of the cellulose
XX synthase complex of Vigna angularis. The invention relates to subunits of
XX cellulose synthetic equipment, that can be used to increase the amount of
XX cellulose synthesized by a plant. The proteins and genes encoding them
XX can also be used to improve the properties of the cellulose being
XX produced by a plant

XX Sequence 10732 BP; 3149 A; 1212 C; 2074 G; 2046 T; 0 U; 2251 Other;

Query Match 5.5%; Score 41.2; DB 3; Length 10732;

Best Local Similarity 14.7%; Pred. No. 0.13; Matches 64; Conservative 182; Mismatches 189; Indels 1; Gaps 1;

QY 262 GTGGGCGCAGCGCCCTGGGCGCATCGCGCTCACCACATTGCTGTGACTTCCGCCAC 321

DB 9949 RYSSSTNCYSYCCYSSRYSTTNSFCRTTYSYSTBTBYSYSRGSYSSRGS 9890

QY 322 GTATTCTATGCGGTTTCATTCGCGCTGATGTCATAAAACCCATGCGCGTTCTAT 381

DB 9889 RAGCYNSTNCYDASTDBYSRCCYTYSSSTDSSTDSSTDSSTDSSTDSSTDSST 9830

QY 382 TCGGTTTGGCGCTTATGACGAAGCCTAGCAGTCACTCGCGCGCGCGCGCGCTGG 441

DB 9829 SDSTSTYRCRSRDYDABSDNSTNCYDASRTBTBSTNCYARCYBYDARCSRDSTYSSR 9770

QY 442 TCGGCGTGGCGACTTCTCATGCAATAGAGTTTCACTCCTACGCGGTATTCGCGGT 501

DB 9769 GIDANSTYSRYSRYSSTYSSTYSSTYSSTYSSTYSSTYSSTYSSTYSSTYSST 9710

QY 502 CTCACCGAGTGGCGAGTGCAGAGTGAATGATCTTTGAAATTAAGGCGCTCG-AGTTGCG 560

DB 9709 YTBYSRRCYCCYDAYSCRYDARCYDACSYSNSTCYDATTSTYSSTYSSTYSSTY 9650

QY 561 CCTTGCTCTCTCTTTGTCAGCGTGAATTGATTCGCGCGAAGAAAGACAGATCCC 620

DB 9649 TTSRCTBT 9590

QY 621 TTCTCTGCTGCTGCGAGGTTTGAAGCTTCAACATGCTCTTGAGTATTCAGGTCAGGC 680

DB 9589 CYSCYTDSSTCYSTTYSSTYAYSCITTSRGSYDASTSTYSRCTTSTYSSTYSDC 9530

QY 681 CCTATTGCGGCGCGCTG 696

DB 9529 SDYSTTBNSTYSDDS 9514

Search completed: April 19, 2004, 11:42:08
Job time : 352.669 secs


```
XX 05-OCT-1999; 99DE-01047792.
XX (DEGS) DEGUSSA-HUELS AG.
XX Moeckel B, Pfeifferle W, Puenhler A, Kalinowski J, Bathe B;
XX WPI; 2001-292927/31.
XX P-PSDB; AAB70881.
XX New lrp gene from coryneform bacteria, used to prepare transformants with
XX increased synthesis of amino acids, particularly lysine and isoleucine.
XX Claim 4; Page 14-15; 22pp; German.
XX This invention describes a novel isolated nucleic acid (I) from
XX coryneform bacteria used for the fermentative production of selected L-
XX amino acids, by fermenting the amino acid-producing coryneform in which
XX at least the lrp gene has been weakened or amplified, then isolating
XX amino acids that have accumulated in the medium or cells. (I) is used to
XX transform coryneforms for production of L-amino acids, specifically
XX lysine and isoleucine, which are used in medicine and particularly as
XX animal feed supplement. It may also be used as probes and primers for
XX isolating related sequences. Regulating expression of (I) improves
XX production of amino acids, especially of L-lysine. This sequence encodes
XX the Corynebacterium glutamicum lrp protein which is used in the method
XX described in the invention
XX
XX Sequence 715 BP; 166 A; 173 C; 178 G; 198 T; 0 U; 0 Other;
SQ
XX
XX Query Match 11.3%; Score 85; DB 4; Length 715;
XX Best Local Similarity 100.0%; Pred. No. 9.6e-16;
XX Matches 85; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GTTCAGAAAAACGGAAGATCTTCAAGCTCGAGGTGCTCCGATCCAGGACAGCCTG 60
DB 85 GTTCAGAAAAACGGAAGATCTTCAAGCTCGAGGTGCTCCGATCCAGGACAGCCTG 26
QY 61 GAACGAGATGATTAAGTTATCGGC 85
DB 25 GAACGAGATGATTAAGTTATCGGC 1
DB
XX
XX RESULT 14
XX ADA71938/c
XX ADA71938 standard; DNA; 2000 BP.
XX
XX ADA71938;
XX
XX 20-NOV-2003 (first entry)
XX
XX Rice gene, SEQ ID 5263.
XX
XX Plant; bacterial infection; fungal infection; viral infection; rice;
XX gene; ds.
XX
XX Oryza sativa.
XX
XX WO2003000898-A1.
XX
XX 03-JAN-2003.
XX
XX 22-JUN-2001; 2001WO-IB001105.
XX
XX 22-JUN-2001; 2001WO-IB001105.
XX
XX (SYGN) SYNGENTA PARTICIPATIONS AG.
XX
XX Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y,
XX Katagiri F, Quan S, Tao Y, Whitman S, Xie Z, Zhu T, Zou G;
XX WPI; 2003-175290/17.
XX
```

```
PT Identifying at least one gene involved in plant resistance or response to
PT pathogenic infection for conferring resistance or tolerance to a plant to
PT bacterial, fungal or viral infection by determining or detecting plant
PT gene expression.
XX
XX Claim 27; SEQ ID NO 5263; 899pp; English.
XX
XX The present invention relates to a method (M1) for identifying genes
XX involved in plant resistance or response to pathogenic infection. M1
XX comprises identifying a gene whose expression is significantly altered in
XX the incompatible interaction of plant gene expression relative to
XX expression of the gene in an uninfected plant, in a mutant plant that
XX does not express a gene associated with response to pathogenic infection,
XX or in a corresponding incompatible or compatible interaction. (M1) is
XX useful for conferring resistance or tolerance to a plant to
XX bacterial, fungal or viral infection. The present sequence was used to
XX illustrate the invention.
XX
XX Sequence 2000 BP; 336 A; 265 C; 284 G; 363 T; 0 U; 752 Other;
SQ
XX
XX Query Match 6.2%; Score 47; DB 7; Length 2000;
XX Best Local Similarity 10.2%; Pred. No. 0.001;
XX Matches 75; Conservative 328; Mismatches 323; Indels 11; Gaps 2;
QY 8 AAACGAGAGATTCATTCAGGCTGAGGTGCGGATTCAGGAGGAGCCGGAACGAG 67
DB 736 RMRMSRWRCSRMKCATKYASARWTKRAKRSYRRRMYRKRKMYRYSRMSCRM 677
QY 68 ATGATTAAGTTATCGGCGGTACGAAATCGCGAAGGTCTAAATCCCTGCTGGCAG 127
DB 676 RARMSKRRKWAAGSAMS CMYWRGASMYSKYSCAKCKKTRWTSMTSGMTGMS 617
QY 128 GTTTGGCATATGACCCGATGATTCGTTGCTCTGTTGTTATTCATACGCTACG 187
DB 616 S--YKSMWSKMSYMGKMTCTMYTSKGSTRSKQMRMSGMRMYRMKRRKRYMR 559
QY 188 AATGATGAGGAGCCGACCTGTTCCGCGCTGATTTCCGGGGCTCCACGGAATGCTG 247
DB 558 YMKRKCCTWRRCMCYRMWGYTWTTSRSRMYTGGYKARYTSKRYMYMKRYMYIYGM 499
QY 248 TCATCGCCCT-----CGTTGGGCGCAGCGCCCTGGCGGCATCGCCTCACCA 298
DB 498 YMKCSYMRKRYGCAKCKCCYACMCWKAAYSMMWYRKYSKXMRMSTKXWMSWYKCR 439
QY 299 CATTGCTGTGAATTCGCGACGATATTCATGCGTTTATTCGCGCTGCATGTGTC 358
DB 438 SMKYGAGCGYCKRMVYCSYGMKMYTMSYKSRCYKRYMYRMYGMVYSAVSS 379
QY 359 AAACCCGATTCGCGCTTTCATTCGCTTTCGCGCTTATCGAGGAGCCTACGACTCA 418
DB 378 MMTYTYIAKMKRYRGRGTMISYKSKYKCTWMCVWKRCYMRMRKRYKXK 319
QY 419 CTGGGCGAGCGCGCAGGCTGTGCGGCTGCGCATTCATTCAGCAATACGCTTC 478
DB 318 RCYCMRYATCYCCYRKGYSRSMRTAGKMKMSWSRMSGYSWYMYKMKMSY 259
QY 479 ACTGCTACTGGGATTCGGGCGCTCAACGAGGAGCGAGATTCATTCCTTTG 538
DB 258 YMSYGMASSGTMRSAAKRTYKYSTSRAMKMRACRMYSACRYRTSYCCSCYGS 199
QY 539 AAATTAAGGCTCGAGTTCGCTTTCCTCTCTTTGTCACGCTGATTTGATTCCT 598
DB 198 SKMKYMSKSCMRWTCSSWCCYCYGAMCWSGMSMMYMGSCGYTRGMKRSYMS 139
QY 599 GCGGAGCAAAAAGACAGATCCCTCTCTGCTGCTCGAGGTTGAGCTTACATGCTC 658
DB 138 CCKKYCSCTTCYCTGTGTCYKMYKSYKCYCYCYWYSYRMYMKKCSGSSMM 79
QY 659 TTGNGTATTCAGAGTCCGCTATTTGGGCGCTGCTGATCTTGGGTGTGGA 718
DB 78 SCAYCSTSTSRMMSMYAAKMGCGSSGMYRMSKSCMYSKYSCKYTGSKCTKRY 19
QY 719 CCATCGGCTACTTCTTC 735
```

Corynebacterium. Corynebacterium bacteria are useful for producing amino acids, nucleic acids, vitamins, saccharides and organic acids, particularly L-lysine. The present sequence is a nucleic acid described in the exemplification of the invention. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from the European Patent Office

Sequence 309400 BP; 70133 A; 86477 C; 83115 G; 69675 T; 0 U; 0 Other;

Query Match 16.9%; Score 127.4; DB 5; Length 309400;
Best Local Similarity 51.5%; Pred. No. 9,66-28;
Matches 318; Conservative 0; Mismatches 296; Indels 3; Gaps 1;

122 CTGACGATTTGGGCGATGATCCGATTTGGCTTGTCTTGTGTTTCAATACG 181
DB 289969 CGGTGGGCTGATTCGGCTGGTGGCTGTGTATGATGACAG 289910
182 GGTACGAATGATGGGAGCCGATGTTTCCGGCTGATTTTGGGGCTCCACCGAA 241
DB 289909 GTTTCGCTGATGATGACGCGATTTTCTTGTGATGATGACGCTTGTGAT 289850
242 TGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 301
DB 289849 TTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 289790
302 TGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 361
DB 289789 TATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 289730
362 ACCCATTTGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 421
DB 289729 CGGGGCGGCGGCGGCTGATTCACCTACGCGCTTACGAGACGATGATGAT 289670
422 CGGCGGCGGCGGCGGCTGATTCACCTACGCGCTTACGAGACGATGATGAT 478
DB 289669 CAGCCCGGCGGCGGCTGATTCACCTACGCGCTTACGAGACGATGATGAT 289610
479 ACTCTACTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 538
DB 289609 AAGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 289550
539 AATTAAGGCGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 598
DB 289549 ATCTAAGGCGCTGATGATGATGATGATGATGATGATGATGATGATGAT 289490
599 GCGAAGCAAAAGAGATCCCTTCTGCTGCTGAGATGATGATGATGATGATGAT 658
DB 289489 TCAAAATTAAGAGATGATGATGATGATGATGATGATGATGATGATGAT 289430
659 TTGTGTAATTCAGATGATGATGATGATGATGATGATGATGATGATGATGAT 718
DB 289429 GTTGTGCGGCGGCGGAGATGATGATGATGATGATGATGATGATGAT 289370
719 CCATCGGATCTCTTC 735
DB 289369 TTCTCGCGCTCCGCTTC 289353

RESULT 12
ACA01968
ID ACA01968 standard; DNA; 177 BP.
ACA01968;
XX
AC 04-JUN-2003 (first entry)
XX
DT C. glutamicum derived ORF SEQ ID 1959.
XX
DE Corynebacterium derived ORF SEQ ID 1959.
XX
KW Corynebacterium glutamicum.
XX
OS Corynebacterium glutamicum.
XX
PN DE10128510-A1.

XX
PD 19-DEC-2002.
XX
PF 13-JUN-2001; 2001DE-01028510.
XX
PR 13-JUN-2001; 2001DE-01028510.

XX
PA (DEBS) DEUSGA AG.
XX
PI Farwick M, Moeckel B, Pfeifferle W, Bathe B, Hutmacher K;
XX
DR WPI; 2003-27970/28.
XX

PT New nucleic acid array useful for monitoring mRNA expression of
PT Corynebacterium glutamicum during fermentation, comprising nucleic acid
PT from Corynebacterium glutamicum.

PS Claim 1; Page 647; 709pp; German.

XX
CC This invention describes a novel nucleic acid array involving
CC Corynebacterium glutamicum polynucleotides. The arrays are used to
CC analyze C. glutamicum, particularly for monitoring a fermentation process
CC to determine expression levels of C. glutamicum cellular mRNA. Such
CC monitoring particularly differentiates between expression levels of
CC different strains of C. glutamicum and allows the adjustment of different
CC culture and fermentation conditions. ACA0001-ACA02188 represent C.
CC glutamicum derived polynucleotides described in the disclosure of the
CC invention

XX
SQ Sequence 177 BP; 54 A; 40 C; 47 G; 36 T; 0 U; 0 Other;

Query Match 14.1%; Score 106; DB 7; Length 177;
Best Local Similarity 100.0%; Pred. No. 1,56-22;
Matches 106; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTGCAAAACCGCAAGATTCATTCAGCTGAGATGTCGCAATCCAGGACGCTG 60
DB 72 GTGCAAAACCGCAAGATTCATTCAGCTGAGATGTCGCAATCCAGGACGCTG 131
QY 61 GAACCATGATTAAGGTTATCGGCGCTACGAATCGCGCAAGTC 106
DB 132 GAACCATGATTAAGGTTATCGGCGCTACGAATCGCGCAAGTC 177

RESULT 13
AAF61688/c
ID AAF61688 standard; DNA; 715 BP.
AAF61688;
XX
AC 12-JUN-2001 (first entry)
XX
DT C. glutamicum lrp encoding DNA.
XX
DE Fermentation; L-amino acid production; lrp gene; lysine; isoleucine;
XX
KW medicine; animal feed supplement; ds.
XX
OS Corynebacterium glutamicum.

XX
FH Key Location/Qualifiers
XX
FT 62..67
XX
FT -35_signal
XX
FT 88..93
XX
FT -10_signal
XX
FT 151..615
XX
FT CDS
XX
FT /*tag= b
XX
FT /*tag= c
XX
FT /product= "lrp"

XX
PD EPI090993-A1.
XX
PN 11-APR-2001.
XX
PF 29-SEP-2000; 2000EP-00121159.

Db 593 GTTTTGGGCGCGAGATGCTGTATCGCTTTGACCAAGTACTTTTGTATCTTC 652
Qy 719 CCAATCGGTAATCTTTC 735
Db 653 TTCTCCGCGTCCGCTTC 669

RESULT 10

ACA01294 standard; DNA; 711 BP.

ACA01294;

03-JUN-2003 (first entry)

C. glutamicum derived ORF SEQ ID 1285.

Coryneform; nucleic acid array; fermentation; culture; ds.

Corynebacterium glutamicum.

DE10128510-A1.

19-DEC-2002.

13-JUN-2001; 2001DB-01028510.

13-JUN-2001; 2001DE-01028510.

(DEGS) DEGUSA AG.

Farwick M, Moeckel B, Pfefferle W, Bathe B, Huthmacher K;

WPI; 2003-279970/28.

New nucleic acid array useful for monitoring mRNA expression of
Corynebacterium glutamicum during fermentation, comprising nucleic acid
from Corynebacterium glutamicum.

Claim 1; Page 456; 709pp; German.

This invention describes a novel nucleic acid array involving
Corynebacterium glutamicum polynucleotides. The arrays are used to
analyze C. glutamicum, particularly for monitoring a fermentation process
to determine expression levels of C. glutamicum cellular mRNA. Such
monitoring particularly differentiates between expression levels of
different strains of C. glutamicum and allows the adjustment of different
culture and fermentation conditions. ACA00010-ACA02188 represent C.
glutamicum derived polynucleotides described in the disclosure of the
invention

Sequence 711 BP; 123 A; 150 C; 198 G; 200 T; 0 U; 0 Other;

Query Match 16.9%; Score 127.4; DB 7; Length 711;
Best Local Similarity 51.5%; Pred. No. 6.3e-29;
Matches 318; Conservative 0; Mismatches 296; Indels 3; Gaps 1;

Qy 122 CTGCAAGTTTGGGATGACCGATTTGATTCGTTGCTCTCTGTTATTCATTCG 181
Db 53 CGGTCCGTTTGGGCTATTCGGCTGGTTGCGCTTTGGCGCTTGTGATGTCAGACG 112
Qy 182 GCTACGAATGTTGGGACCGCCACTGTTTCCGCGCTGATTTTCGGGGCTCCACGAAA 241
Db 113 GTTTCGCTGGTGTGACGCGCATTTCTCTTCGATCTATGCGGTTGATGAGAAAT 172
Qy 242 TGCTGTATCGCCCTGTGTGTGTGGCGACGCGCCCTGGCGCGCATCGGCTCACCACAT 301
Db 173 TTCTGGGAATCGGATGTCACCGCAGGATTCGCGCCGTTTGGCGCGCTGGT 232
Qy 302 TGCGTGAATTCGCGCAGTATTCATTCATTCGTTCCCGCGCTGATGTCAGTCAAA 361
Db 233 TCATGTGAATTTCCGCAATTTTCTACGCTCTACCTTCCACGCCACCGCATCAAGT 292

Qy 362 ACCCATTCGCGCTTCTATTCGCTTTTCGCGCTTATGACGAAGCTTACGAGTCACTG 421
Db 293 CCGGCGCGCGCGCGCGCTATTCACCTACGCGCTTACGAGAGATCTACGCCATCGT 352
Qy 422 CGGCGAGCGCGCGAGCTGTGCGCTG---GCACTTATCTCAATGCAATAGCGTTTC 478
Db 353 CAGCGCGCGCGCGCTGCGATATCAGTGCACGCGGCTCTTACCGTTCAATTTTGGCC 412
Qy 479 ACTCTACTGAGTATTCGCGCTGCTACCGAGTGCAGATGAGTGAATTCCTTTTG 536
Db 413 AAGCTGTGCTGTATCCAGGAATTTATGCGCTTGTGTGTCAATAGTCTGCCATG 472
Qy 539 AATTGAAGGCTTCGAGTTGCGCTTTGCTCTCTTTGTACGCTGACTTGAATTCCT 598
Db 473 ATCTAAAGCATGTGATTTTGGCTGACCGCGCTTTGTGTGTGCGTGGAGGAGAT 532
Qy 599 GCCGACGAAAAAGCATCCCTCTCTGCTGCTGCGAGTTTGAAGTTGACCAATGCTC 658
Db 533 TCAAAATTAACAAGATTTTGTGCTGCAATTTTGCCTGTGATGCTGTGCTGCTGCTG 592
Qy 659 TTGTGTAATTCAGGTGAGCCCTATTTTGGCGCGCTGCTGATCTTTGGGCTCTTTGA 718
Db 593 GTTTTGGCGCGCGAGCATGCTGTATGCTTTGACCAAGTATTTGATCTTC 652
Qy 719 CCAATCGGTAATCTTTC 735
Db 653 TTCTCCGCGTCCGCTTC 669

RESULT 11

AAH68534/c
AAH68534 standard; DNA; 309400 BP.

AAH68534;

26-SEP-2001 (first entry)

C glutamicum coding sequence fragment SEQ ID NO: 7069.

Coryneform bacterium; amino acid synthesis; vitamin; saccharide;

organic acid synthase; ds.

Corynebacterium glutamicum.

EPI108790-A2.

20-JUN-2001.

18-DEC-2000; 2000EP-00127688.

16-DEC-1999; 99UP-00377484.

07-APR-2000; 2000UP-00159162.

03-AUG-2000; 2000UP-00280988.

(KYOW) KYOWA HAKKO KOGYO KK.

Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;

Tateishi N, Senoh A, Ikeda M, Ozaki A;

WPI; 2001-376931/40.

Novel polynucleotides derived from Coryneform bacteria, for identifying

mutation point of a gene, measuring expression of a gene, analyzing

expression profile or pattern of a gene and identifying homologous gene.

Disclosure; SEQ ID NO 7069; 246pp + Sequence Listing; English.

The present invention provides a number of nucleotide and protein

sequences from the Coryneform bacterium Corynebacterium glutamicum. These

are useful for identifying the mutation point of a gene derived from a

mutant of coryneform bacterium, measuring expression amount and analyzing

the expression profile or expression pattern of a gene derived from

Coryneform bacterium, and identifying a homologue of a gene derived from

XX	Moeckel B, Pfefferle W, Fuehrer A, Kalinowski J, Bathe B;
PI	
XX	WPI; 2001-292927/31.
DR	
XX	New lrp gene from corynebacter bacteria, used to prepare transformants with
PT	increased synthesis of amino acids, particularly lysine and isoleucine.
XX	
PS	Example 3; Page 17; 22pp; German.
XX	
CC	This invention describes a novel isolated nucleic acid (I) from
CC	coryneform bacteria used for the fermentative production of selected L-
CC	amino acids, by fermenting the amino acid-producing coryneform in which
CC	at least the lrp gene has been weakened or amplified, then isolating
CC	amino acids that have accumulated in the medium or cells. (I) is used to
CC	transform coryneforms for production of L-amino acids, specifically
CC	lysine and isoleucine, which are used in medicine and particularly as
CC	animal feed supplement. It may also be used as probes and primers for
CC	isolating related sequences. Regulating expression of (I) improves
CC	production of amino acids, especially of L-lysine. This sequence encodes
CC	a fragment of the Corynebacterium glutamicum lrp protein which is used in
CC	the method described in the invention
SQ	
	Sequence 778 BP; 175 A; 208 C; 207 G; 188 T; 0 U; 0 Other;
	Query Match 36.3%; Score 273.6; DB 4; Length 778;
	Best Local Similarity 98.6%; Pred. No. 2.3e-74;
	Matches 276; Conservative 0; Mismatches 4; Indels 0; Gaps 0
OY	1 GTGCAAAAAACCGAAGATTCATTCAACCCTGGAGGTGTGCCCATCCAAAGCAGCCCTG 60
Dd	284 GTGCAAAAACCGAAGAATTCATTCAACCCTGGAGGTGTGCCCATCCAAAGCAGCCCTG 225
OY	61 GAACCAAGATGAAGAAGTATGCGGCCTACGAAATGCGCGCAAGTCTAAAACCTCCCT 120
Dd	224 GAACCAAGATGAAGAAGTATGCGGCCTACGAAATGCGCGCAAGTCTAAAACCTCCCT 165
OY	121 GCTGCAGTTTGCGGCATGTACCCGATGTGTATGCGTTTGCTCTTGTTATTCANIC 180
Dd	164 GCTGCAGTTTGCGGCATGTACCCGATGTGTATGCGTTTGCTCTTGTTATTCANIC 105
OY	181 GGCTACGATGTGTGGGCGACCCCACACTGTTTCCGGCTGATTTTGGCGGGCTCCACCGA 240
Dd	104 GGCTACGATGTGTGGGCGACCCCACACTGTTTCCGGCTGATTTTGGCGGGCTCCACCGA 45
OY	241 ATGCTGTGATCGCCCTGCTGTGTGGCGGACGCGCCCCCTG3 280
Dd	44 ATGCTGTGATCGCCCTGCTGTGTGGCGGACGCGCCCCCTG3 5
	RESULT 9
	AAH68373
ID	AAH68373 standard; DNA; 711 BP.
XX	
AC	AAH68373;
DT	26-SEP-2001 (first entry)
XX	
C	C glutamicum coding sequence fragment SEQ ID NO: 3408.
DE	
XX	Coryneform bacterium; amino acid synthesis; vitamin; saccharide;
KM	organic acid synthetase; ds.
XX	
OS	Corynebacterium glutamicum.
XX	
PN	BPI108790-A2.
XX	
PD	20-JUN-2001.
PF	18-DEC-2000; 2000EP-00127688.
XX	
PR	16-DEC-1999; 99JP-00377484.
PR	07-APR-2000; 2000JP-00159162.

Query	Seq	Sequence	711 BP, 123 A, 190 C, 198 G, 200 T, 0 U, 0 Other;
XX	03 -AUG-2000; 2000JP-00280988.		
PR	(KXOM) KYOMA HAKKO KOGYO KK.		
PA	Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;		
PI	Tateishi N, Senoh A, Ikeda M, Ozaki A;		
XX	WPI: 2001-376931/40.		
DR	P-PSDB: AAG93154.		
XX	Novel polynucleotides derived from Corynebacterium bacteria, for identifying		
PT	mutation point of a gene, measuring expression of a gene, analyzing		
PR	expression profile or pattern of a gene and identifying homologous gene.		
PS	Claim 8; SEQ ID NO 3408; 246pp + Sequence Listing; English.		
XX	The present invention provides a number of nucleotide and protein		
CC	sequences from the Corynebacterium bacterium Corynebacterium glutamicum. These		
CC	are useful for identifying the mutation point of a gene derived from a		
CC	mutant of corynebacterium bacterium, measuring expression amount and analyzing		
CC	the expression profile or expression pattern of a gene derived from		
CC	Corynebacterium bacterium, and identifying a homologue of a gene derived from		
CC	Corynebacterium bacterium. Corynebacterium bacteria are useful for producing amino		
CC	acids, nucleic acids, vitamins, saccharides and organic acids,		
CC	particularly L-lysine. The present sequence is a nucleic acid described		
CC	in the exemplification of the invention. Note: The sequence data for this		
CC	patent did not form part of the printed specification, but was obtained		
CC	in electronic format directly from the European Patent Office		
XX	Sequence 711 BP, 123 A, 190 C, 198 G, 200 T, 0 U, 0 Other;		
Seq	Query Match	16.9%; Score 127.4; DB 5; Length 711;	
	Best Local Similarity	51.5%; Pred. No. 6,3e-29;	
	Matches 318; Conservative	0; Mismatches 296; Indels 3; Gaps 1	
Qy	122	CTGCAGGTTGGGCGATGACCGGATGGATTTGGTTGGTGTCTGTGGTATCAATACG	181
Db	53	CGGTGGTTGGGCGCTGATTCGCTGGTTGGCTTTGGGCTGTGATGATGTCAGAACG	112
Qy	182	GCTACGAATGGTGGCGACGCCCATGTTTTCGCGGCTGATTTTCGCGGCTCCACCGAAA	241
Db	113	GTTTGCGCTGGTGGTGAAGCGCGGATTTTCTCTTGATCATACCGGATGGAAT	172
Qy	242	TGCTGTCATGCGCTCGTTGGTGGCGGACGCGCCCTGGGGCGCATCGCGCTACACAT	301
Db	173	TTCTGGCAATCGGATGTCACCGGATTCGCGCGGTTTGGCGGGGTGGTGGT	232
Qy	302	TGCTGATGAACTTGGCGACGATTCATGCGTTTCAATCCCGCTGATGTCATGATCAAA	361
Db	233	TATGATGAATTTCCGCCAATTTCTACGTCATCCTCCACGACACGCAATCAAGT	292
Qy	362	ACCCGATGCGCGTTTCTAATTCGTTTCGCGTTATGACGAAAGCTTACGATGACTG	421
Db	293	CCGGCGCGCGCGCGCCTAATTCACCTTACGCGCTTACGACAGTCCTAACGCAATCGAT	352
Qy	422	CGGCGAGCGCGCGAGCTGATGGGGTG--GCGACTTATCCATGCAATGCAATAGCGTTTC	478
Db	353	CAGCCCGCCACCTGCGATATCAATGAGCAGCGGGGTGTTACGTTCAATTTGTGTC	412
Qy	479	ACTCTACTGGGTATTTGCGCGGTCTCACCGGATGCGCATGCGCAAGTTGATTCCTTTTG	538
Db	413	AGGCTCTGGGTATTCACGAAATTAATGGGCGCTTGTTGTCAGAGCTGCGCCGATG	472
Qy	539	AAATTAAGGCTCGAGTTGCGCCTTGTCTCTCTTTGTCAGCGTGAATTTGATTCCT	598
Db	473	ATCTAAAGGATGATTTTGGCTCGACCGCGCTGTGTGTGTGTGCGGTGAGAGCAT	532
Qy	599	GCGGAACGAAAGAGATCCCTTCTCTGCTGCTCGCAGATTTGAGCTTACCATTTGCTC	658
Db	533	TCAAAATTAACAAGATTATTCGTCGCATTAATTCGCGGTGATTTGGCTGTGTTCCG	592
Qy	659	TTGTGTAATTCAGGTCAGGCCCTAATTTGGCGCGCTGTGATCTTTCTTGGGTCTGTTA	718

XX Sequence 564 BP; 91 A; 167 C; 144 G; 162 T; 0 U; 0 Other;
SQ
Query Match 73.4%; Score 553; DB 7; Length 564;
Best Local Similarity 99.8%; Pred. No. 2.7e-161;
Matches 564; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
QY 189 ATGTGGGAGGCGCCCACTGTTTTCGGGCTGATTTTCGGGCTCCACGAAATGCTGCT 248
DB 1 ATGTGGGAGGCGCCCACTGTTTTCGGGCTGATTTTCGGGCTCCACGAAATGCTGCT 60
QY 249 CATGCGCCCTGCTGTTGGGAGGAGCGCCCTGAGCGCCATGCGGCTCACCACATTTGCTGCT 308
DB 61 CATGCGCCCTGCTGTTGGGAGGAGCGCCCTGAGCGCCATGCGGCTCACCACATTTGCTGCT 119
QY 309 GAACCTCCGCGCAGTATTTCTATGCGTTTTCATTCGCGCTGATGTTGCTCAAAAACCCCAT 368
DB 120 GAACCTCCGCGCAGTATTTCTATGCGTTTTCATTCGCGCTGATGTTGCTCAAAAACCCCAT 179
QY 369 TGGCGCTTCTATTTGGGTTTGGGCTTATGAGAGAGCTTACGAGTCACTGCGGCGCAG 428
DB 180 TGGCGCTTCTATTTGGGTTTGGGCTTATGAGAGAGCTTACGAGTCACTGCGGCGCAG 239
QY 429 GCGCGAGGCTGCTGCTGCGCTGAGGCTTATCTCAATGCAATAGCGTTTCACTCTACTG 488
DB 240 GCGCGAGGCTGCTGCTGCGCTGAGGCTTATCTCAATGCAATAGCGTTTCACTCTACTG 299
QY 489 GGTATTCGCGGCTTCAACCGGAGTGGCGATTCGCAAGTGGATTCCTTTTGAATTAAGG 548
DB 300 GGTATTCGCGGCTTCAACCGGAGTGGCGATTCGCAAGTGGATTCCTTTTGAATTAAGG 359
QY 549 CCTGAGTTGCGCCCTTGTCTCTCTTTGTCAAGCTGATTTGATTCCTGCGGAGCA 608
DB 360 CCTGAGTTGCGCCCTTGTCTCTCTTTGTCAAGCTGATTTGATTCCTGCGGAGCA 419
QY 609 AAGGAGATCCCTTCTCTGCTGCTGCGAGTTTGAAGCTTCAACATTTGCTTGTAT 668
DB 420 AAGGAGATCCCTTCTCTGCTGCTGCGAGTTTGAAGCTTCAACATTTGCTTGTAT 479
QY 669 TCCAGGTCAGGCGCTTATTTGCGGCGCTGCTGATTTCTTGGTCTGTTGACCATCGGTA 728
DB 480 TCCAGGTCAGGCGCTTATTTGCGGCGCTGCTGATTTCTTGGTCTGTTGACCATCGGTA 539
QY 729 CTTCCTTCTTGGGAAAGGCTGCTAAA 753
DB 540 CTTCCTTCTTGGGAAAGGCTGCTAAA 564
RESULT 7
AAF61693/c
ID AAF61693 standard; DNA; 397 BP.
XX
AC AAF61693;
XX
DT 12-JUL-2001 (first entry)
XX
DE C. glutamicum lrp encoding DNA fragment SEQ ID 7.
XX
KM Fermentation; L-amino acid production; lrp gene; lysine; isoleucine;
KW medicine; animal feed supplement; ds.
XX
OS Corynebacterium glutamicum.
XX
PN EP1090993-A1.
XX
PD 11-APR-2001.
XX
PF 29-SEP-2000; 2000EP-00121159.
XX
PR 05-OCT-1999; 99DE-01047792.
XX
PA (DEGUS) DEGUS-A-HUELS AG.
XX

PI Moeckel B, Pfeifferle W, Puehler A, Kalinowski J, Bathe B;
XX WPI, 2001-292927/31.
DR
XX New lrp gene from corynebacteria, used to prepare transformants with
PT increased synthesis of amino acids, particularly lysine and isoleucine.
XX
PS Example 3; Page 16; 22pp; German.
XX
CC This invention describes a novel isolated nucleic acid (I) from
CC corynebacteria used for the fermentative production of selected L-
CC amino acids, by fermenting the amino acid-producing corynebacteria in which
CC at least the lrp gene has been weakened or amplified, then isolating
CC amino acids that have accumulated in the medium or cells. (I) is used to
CC transform corynebacteria for production of L-amino acids, specifically
CC lysine and isoleucine, which are used in medicine and particularly as
CC animal feed supplement. It may also be used as probes and primers for
CC isolating related sequences. Regulating expression of (I) improves
CC production of amino acids, especially of L-lysine. This sequence encodes
CC a fragment of the Corynebacterium glutamicum lrp protein which is used in
CC the method described in the invention
XX
SQ Sequence 397 BP; 96 A; 100 C; 104 G; 97 T; 0 U; 0 Other;
Query Match 36.3%; Score 273.6; DB 4; Length 397;
Best Local Similarity 98.6%; Pred. No. 1.7e-74;
Matches 276; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 GTGGAAAAAGCAAGATTCATTTCAAGCTGAGAGTGGCGCATCCAGGAGGCGCTG 60
DB 284 GTGGAAAAAGCAAGATTCATTTCAAGCTGAGAGTGGCGCATCCAGGAGGCGCTG 225
QY 61 GAACCAATGATTAAGTTATTCGCGCTACGAATGCGCAGGCTTAAACCTCCCTT 120
DB 224 GAACCAATGATTAAGTTATTCGCGCTACGAATGCGCAGGCTTAAACCTCCCTT 165
QY 121 GCTGAGCTTTGGGCAATGACCGATTTGGATGCGTTGCTCTTGGTATTCATATC 180
DB 164 GCTGAGCTTTGGGCAATGACCGATTTGGATGCGTTGCTCTTGGTATTCATATC 105
QY 181 GGTACGAATGTTGGGAGGCGCCCACTGTTTTCGCGCTGATTTTTCGCGGCTCCACCGAA 240
DB 104 GGTACGAATGTTGGGAGGCGCCCACTGTTTTCGCGCTGATTTTTCGCGGCTCCACCGAA 45
QY 241 ATGTGCTCATCGCCCTCTGTTTGGGAGGCGGCGCCCTCTG 280
DB 44 ATGTGCTCATCGCCCTCTGTTTGGGAGGCGGCGCCCTCTG 5
RESULT 8
AAF61695/c
ID AAF61695 standard; DNA; 778 BP.
XX
AC AAF61695;
XX
DT 12-JUL-2001 (first entry)
XX
DE C. glutamicum lrp encoding DNA fragment SEQ ID 9.
XX
KM Fermentation; L-amino acid production; lrp gene; lysine; isoleucine;
KW medicine; animal feed supplement; ds.
XX
OS Corynebacterium glutamicum.
XX
PN EP1090993-A1.
XX
PD 11-APR-2001.
XX
PF 29-SEP-2000; 2000EP-00121159.
XX
PR 05-OCT-1999; 99DE-01047792.
XX
PA (DEGUS) DEGUS-A-HUELS AG.
XX

XX C glutamicum coding sequence fragment SEQ ID NO: 1.
XX
XX Corynebacterium; amino acid synthesis; vitamin; saccharide;
XX organic acid synthesis; ds.
XX Corynebacterium glutamicum.
XX
XX Epi108790-A2.
XX
XX 20-JUN-2001.
XX
XX 18-DEC-2000; 2000EP-00127688.
XX
XX 16-DEC-1999; 99JP-00377484.
XX 07-APR-2000; 2000JP-00159162.
XX 03-AUG-2000; 2000JP-00280988.
XX
XX (KYOM) KYOMA HAKKO KOGYO KK.
XX
XX Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;
XX Tateishi N, Senoh A, Ikeda W, Ozaki A;
XX WPI; 2001-376931/40.
XX
XX Novel polynucleotides derived from Corynebacterium bacteria, for identifying
XX mutation point of a gene, measuring expression of a gene, analyzing
XX expression profile or pattern of a gene and identifying homologous gene.
XX
XX
XX Claim 7, SEQ ID NO 1; 246pp + Sequence listing; English.
XX
XX The present invention provides a number of nucleotide and protein
XX sequences from the Corynebacterium glutamicum. These
XX are useful for identifying the mutation point of a gene derived from a
XX mutant of corynebacterium, measuring expression amount and analyzing
XX the expression profile or expression pattern of a gene derived from
XX Corynebacterium, and identifying a homologue of a gene derived from
XX Corynebacterium. Corynebacterium bacteria are useful for producing amino
XX acids, nucleic acids, vitamins, saccharides and organic acids,
XX particularly L-lysine. The present sequence is a nucleic acid described
XX in the exemplification of the invention. Note: The sequence data for this
XX patent did not form part of the printed specification, but was obtained
XX in electronic format directly from the European Patent Office
XX
XX
XX Sequence 349980 BP; 79703 A; 91547 C; 98381 G; 80349 T; 0 U; 0 Other;
XX
XX Query Match 100.0%; Score 753; DB 5; Length 349980;
XX Best Local Similarity 100.0%; Pred. No. 3,1e-222;
XX Matches 753; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX
XX 1 GTGCAAAAAAGCAGAGATTATTCAGCTTGAGGTGTGCGCATCCAGGACGCTTG 60
XX
XX 276829 GTGCAAAAAAGCAGAGATTATTCAGCTTGAGGTGTGCGCATCCAGGACGCTTG 276888
XX
XX 61 GAACGAGATGATTAAGTTATCGGCGCTACGAAATCGGCGAGGCTTAAAACTCCCTT 120
XX
XX 276889 GAACGAGATGATTAAGTTATCGGCGCTACGAAATCGGCGAGGCTTAAAACTCCCTT 276948
XX
XX 121 GCTGAGGTTTGGGCGATGACCGATGTTGCGTTTGTCTCTTGGTTATTCATAC 180
XX
XX 276949 GCTGAGGTTTGGGCGATGACCGATGTTGCGTTTGTCTCTTGGTTATTCATAC 277008
XX
XX 181 GGCTAGAGATGTTGGGCGACGCCACGTTTCCGGCTGATTTTGGCGGGCTCCACCGAA 240
XX
XX 277009 GGCTAGAGATGTTGGGCGACGCCACGTTTCCGGCTGATTTTGGCGGGCTCCACCGAA 277068
XX
XX 241 ATGCTGATGATGCGCTCGTTTGGGCGAGCGCCCTTGGGCGGCGCATGCGCTCACCA 300
XX
XX 277069 ATGCTGATGATGCGCTCGTTTGGGCGAGCGCCCTTGGGCGGCGCATGCGCTCACCA 277128
XX
XX 301 TTGCTGATGATGCGCGACGATTCATGCTTTCATTCCTCGCTGATGATGATGATGAT 360
XX
XX 277129 TTGCTGATGATGCGCGACGATTCATGCTTTCATTCCTCGCTGATGATGATGATGAT 277188

QY 361 AACCCATGTCGCCGTTTCTATTGCGTTTTCGCGCTTATGACGAAAGCTTACGAGTCACT 420
DB 277189 AACCCATGTCGCCGTTTCTATTGCGTTTTCGCGCTTATGACGAAAGCTTACGAGTCACT 277248
QY 421 GCGGCGAGGCGCGGCGGCTGCTGCGGCTGGGCGGCTTATGCAATGCAATAGGCTTTCAC 480
DB 277249 GCGGCGAGGCGCGGCGGCTGCTGCGGCTGGGCGGCTTATGCAATGCAATAGGCTTTCAC 277308
QY 481 TCCTACTGGGTTATTCGCGGCTGCTCACCGGAGTGGCGATGCAAGATTGCTTTTGA 540
DB 277309 TCCTACTGGGTTATTCGCGGCTGCTCACCGGAGTGGCGATGCAAGATTGCTTTTGA 277368
QY 541 ATTAAGGCGCTGAGTTCGCCCTTGTCTCTCTTGTGACGCTGACTTGTGATTCCTG 600
DB 277369 ATTAAGGCGCTGAGTTCGCCCTTGTCTCTCTTGTGACGCTGACTTGTGATTCCTG 277428
QY 601 CGAACGAAAAAGCAGATCCCTTCTCTGCTGCTGCGAGTTTGAAGTTCAACATGCTCTT 660
DB 277429 CGAACGAAAAAGCAGATCCCTTCTCTGCTGCTGCGAGTTTGAAGTTCAACATGCTCTT 277488
QY 661 GTGTAATTCAGATGAGGCTTATTTGGGCGGCTGATCTTGTGGTCTGTTGAC 720
DB 277489 GTGTAATTCAGATGAGGCTTATTTGGGCGGCTGATCTTGTGGTCTGTTGAC 277548
QY 721 ATCCGTAATTCCTTCTTGGGAAAGCTGCTTAA 753
DB 277549 ATCCGTAATTCCTTCTTGGGAAAGCTGCTTAA 277581

RESULT 6
ID ACA01969 standard; DNA; 564 BP.
XX
XX ACA01969;
XX
XX 04-JUN-2003 (first entry)
XX
XX C. glutamicum derived ORF SEQ ID 1960.
XX
XX Corynebacterium; nucleic acid array; fermentation; culture; ds.
XX
XX Corynebacterium glutamicum.
XX
XX DE10128510-A1.
XX
XX 19-DEC-2002.
XX
XX 13-JUN-2001; 2001DE-01028510.
XX
XX 13-JUN-2001; 2001DE-01028510.
XX
XX (DEGS) DEGUSA AG.
XX
XX Farwick M, Moeckel B, Pfefferle W, Bathe B, Huthmacher K;
XX WPI; 2003-279970/28.
XX
XX
XX New nucleic acid array useful for monitoring mRNA expression of
XX Corynebacterium glutamicum during fermentation, comprising nucleic acid
XX from Corynebacterium glutamicum.
XX
XX
XX Claim 1; Page 647; 709pp; German.
XX
XX This invention describes a novel nucleic acid array involving
XX Corynebacterium glutamicum polynucleotides. The arrays are used to
XX analyze C. glutamicum, particularly for monitoring a fermentation process
XX to determine expression levels of C. glutamicum cellular mRNA. Such
XX monitoring particularly differentiates between expression levels of
XX different strains of C. glutamicum and allows the adjustment of different
XX culture and fermentation conditions. ACA0010-ACA02188 represent C.
XX glutamicum derived polynucleotides described in the disclosure of the
XX invention

DB 641 ATTAAGGCGCTCGAGTTGCGCCCTTTCCTCTCTTTGTCACGCTGACGTTGATTCCTGC 700
QY 601 CGAACGAAAAAGCAGATCCCTTCTCTGCTGCTCGCAGGTTGACCTTACCATTTCTCTT 660
DB 701 CGAACGAAAAAGCAGATCCCTTCTCTGCTGCTCGCAGGTTGACCTTACCATTTCTCTT 760
QY 661 GTGGAATTCCAGGTCAGGCTCCTATTGCGGCGCTGATCTCTTGGGCTGTGACC 720
DB 761 GTGGAATTCCAGGTCAGGCTCCTATTGCGGCGCTGATCTCTTGGGCTGTGACC 820
QY 721 ATCCGTAATCTTCTTCTTGGGAAAGGCTGCTAAA 753
DB 821 ATCCGTAATCTTCTTCTTGGGAAAGGCTGCTAAA 853

RESULT 4

AAH21112 standard; DNA; 1271 BP.

AAH21112;

05-SEP-2001 (first entry)

C. glutamicum DNA encoding brnF and brnE.

XX L-amino acid production; brnF; brnE; branched-chain amino acid;
XX coryneform bacterium; leucine; isoleucine; valine; medicine;
KW animal nutrition; ds.

XX Corynebacterium glutamicum.

XX Key Location/Qualifiers

XX CDS 101..856

XX FT /*tag= a

XX FT /product= "brnF"

XX FT 853..1179

XX FT /*tag= b

XX FT /product= "brnE"

XX EP1096010-A1.

XX 02-MAY-2001.

XX 11-OCT-2000; 2000EP-00122057.

XX 27-OCT-1999; 99DE-01051708.

XX (DEGS) DEGUSA AG.

XX (KERJ) FORSCHUNGSZENTRUM JUELICH GMBH.

XX Kennerknecht N. Eggeling L. Sahm H. Pfeifferle W.

XX MPI; 2001-391595/42.

XX New export genes from coryneform bacteria, useful for increasing

XX fermentative production of branched-chain amino acids.

XX Claim 4 (i); Page 17-18; 23pp; German.

CC This invention describes a novel isolated polynucleotide (I) containing
CC at least one sequence that (i) is 70% identical with a sequence that
CC encodes at least one of 251 amino acid (aa) (3) or 108 aa (5)
CC polypeptides; (ii) encodes a polypeptide at least 70% identical with (3)
CC or (5); (iii) is the complement of (i) or (ii), or (iv) contains at least
CC 15 consecutive bases from (i)-(iii). The invention also describes (a)
CC protein derived from sequences of 753 bp (2) or 324 bp (4); (b)
CC coryneform microorganisms, especially Corynebacterium, transformed with
CC one or more (i), where these are replicative DNA; (c) production of
CC branched-chain L-aa by fermentation of coryneform bacteria in which the
CC brnE and/or brnF genes (or equivalent sequences) are amplified,
CC especially overexpressed; and (d) method for isolating the brnE and/or
CC brnF genes. (i) is used for transformation of coryneform bacteria being
CC used for fermentative production of branched-chain amino acids,

CC specifically leucine, isoleucine and valine, which are useful in medicine
CC and animal nutrition. (i) can also be used as source of primers and
CC probes for isolation of related sequences. Transformation with (i)
CC increases yield of branched-chain amino acids. This sequence encodes the
CC Corynebacterium glutamicum ATCC 13032 brnF and brnE proteins described in
CC the method of the invention

SQ Sequence 1271 BP; 265 A; 328 C; 318 G; 360 T; 0 U; 0 Other;

Query Match 100.0%; Score 753; DB 4; Length 1271;

Best Local Similarity 100.0%; Pred. No. 2,5e-223; Indels 0; Gaps 0;

Matches 753; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTGCAAAAAACGACAGATTCATTCAGCTGAGGTGTGCCATCCAGGCGAGCCCTG 60
DB 101 GTGCAAAAAACGACAGATTCATTCAGCTGAGGTGTGCCATCCAGGCGAGCCCTG 160
QY 61 GAACCGATGATTAAGGTTATGCGCGCTACGAAATGCGCAGAGTCTAAAACCTGCCCT 120
DB 161 GAACCGATGATTAAGGTTATGCGCGCTACGAAATGCGCAGAGTCTAAAACCTGCCCT 220
QY 121 GCTGCAAGTTTGGGCAATGACCGATTTGATTTGCTTGTGTTATTCATATAC 180
DB 221 GCTGCAAGTTTGGGCAATGACCGATTTGATTTGCTTGTGTTATTCATATAC 280
QY 181 GGCTACGATGTGGGCAAGCCCACTGTTTTCGGGCTGATTTTCGGGCGCTCCACCGAA 240
DB 281 GGCTACGATGTGGGCAAGCCCACTGTTTTCGGGCTGATTTTCGGGCGCTCCACCGAA 340
QY 241 ATGCTGTCATGCGCTCGTGTGTTGCGGCGAGCGCCCTGCGGCGCATCCGCTACACAA 300
DB 341 ATGCTGTCATGCGCTCGTGTGTTGCGGCGAGCGCCCTGCGGCGCATCCGCTACACAA 400
QY 301 TTGCTGTGAATCTTCCGCAAGTATTCATTCATTCATTCATTCATTCATTCATTCATTC 360
DB 401 TTGCTGTGAATCTTCCGCAAGTATTCATTCATTCATTCATTCATTCATTCATTCATTC 460
QY 361 AACCCCATTTGCGGCTTCTATTCATTCATTCATTCATTCATTCATTCATTCATTCATTC 420
DB 461 AACCCCATTTGCGGCTTCTATTCATTCATTCATTCATTCATTCATTCATTCATTCATTC 520
QY 421 GCGGCCAGGCGCGAGGCTGTGCGCGTGGCGCACTTATTCATTCATTCATTCATTCATTC 480
DB 521 GCGGCCAGGCGCGAGGCTGTGCGCGTGGCGCACTTATTCATTCATTCATTCATTCATTC 580
QY 481 TCCACTGAGGTTATTCGCGGCTGTACCGGAGTGGCGCATTCGAGATTCCTTTTGA 540
DB 581 TCCACTGAGGTTATTCGCGGCTGTACCGGAGTGGCGCATTCGAGATTCCTTTTGA 640
QY 541 ATTAAGGCGCTGAGATTCGCGCTTGTCTCTCTTGTGACGCTGATTCCTTCTGC 600
DB 641 ATTAAGGCGCTGAGATTCGCGCTTGTCTCTCTTGTGACGCTGATTCCTTCTGC 700
QY 601 CGAACGAAAAAGCAGATCCCTTCTCTGCTGCTCGCAGGTTGACCTTACCATTTCTCTT 660
DB 701 CGAACGAAAAAGCAGATCCCTTCTCTGCTGCTCGCAGGTTGACCTTACCATTTCTCTT 760
QY 661 GTGGAATTCCAGGTCAGGCTCCTATTGCGGCGCTGATCTCTTGGGCTGTGACC 720
DB 761 GTGGAATTCCAGGTCAGGCTCCTATTGCGGCGCTGATCTCTTGGGCTGTGACC 820
QY 721 ATCCGTAATCTTCTTCTTGGGAAAGGCTGCTAAA 753
DB 821 ATCCGTAATCTTCTTCTTGGGAAAGGCTGCTAAA 853

RESULT 5

AAH64966 standard; DNA; 349980 BP.

AAH64966;

26-SEP-2001 (first entry)

```

DB 241 ATGCTGATCAGCCCTGCTGTTGGGCGACGCGCCCTGGGCGGCATCGCGCTCACCA 300
QY 301 TTGCTGATGAATTCGCGCAGATATCTATGCGTTTTCATTCGCGTGCATGTGTCAAA 360
DB 301 TTGCTGATGAATTCGCGCAGATATCTATGCGTTTTCATTCGCGTGCATGTGTCAAA 360
QY 361 AACCCCATGCGCGTTTCTATTCGCGTTTTCGCGTTTTCGCGTTTTCGCGTTTTCGCG 420
DB 361 AACCCCATGCGCGTTTCTATTCGCGTTTTCGCGTTTTCGCGTTTTCGCGTTTTCGCG 420
QY 421 GCGGCGAGCGCGCGAGCGTGTGCGCGTGTGCGCGTGTGCGCGTGTGCGCGTGTGCGCG 480
DB 421 GCGGCGAGCGCGCGAGCGTGTGCGCGTGTGCGCGTGTGCGCGTGTGCGCGTGTGCGCG 480
QY 481 TCCTACTGCGTATTCGCGCGTGTGCGCGTGTGCGCGTGTGCGCGTGTGCGCGTGTGCG 540
DB 481 TCCTACTGCGTATTCGCGCGTGTGCGCGTGTGCGCGTGTGCGCGTGTGCGCGTGTGCG 540
QY 541 ATTAAGGCGCTCGAGTTCGCGCGTGTGCGCGTGTGCGCGTGTGCGCGTGTGCGCGTGTG 600
DB 541 ATTAAGGCGCTCGAGTTCGCGCGTGTGCGCGTGTGCGCGTGTGCGCGTGTGCGCGTGTG 600
QY 601 CGAAGGAAAGAGAGATCCCTTCTCTGCTGCTGCGAGTTCGCGCGTGTGCGCGTGTGCG 660
DB 601 CGAAGGAAAGAGAGATCCCTTCTCTGCTGCTGCGAGTTCGCGCGTGTGCGCGTGTGCG 660
QY 661 GTGTAATTCAGGCGAGCGCGCTATTTGCGGCGTGTGCGCGTGTGCGCGTGTGCGCGTGT 720
DB 661 GTGTAATTCAGGCGAGCGCGCTATTTGCGGCGTGTGCGCGTGTGCGCGTGTGCGCGTGT 720
QY 721 ATCCGTAATTCCTTCTGCGGAAAGCGCTGCTAA 753
DB 721 ATCCGTAATTCCTTCTGCGGAAAGCGCTGCTAA 753
QY 721 ATCCGTAATTCCTTCTGCGGAAAGCGCTGCTAA 753
DB 721 ATCCGTAATTCCTTCTGCGGAAAGCGCTGCTAA 753

RESULT 3
AAH21109
ID AAH21109 standard; DNA; 1271 BP.
AC AAH21109;
DT 05-SEP-2001 (first entry)
DE C. glutamicum DNA encoding brnF and brnE.
XX L-amino acid production; brnF; brnE; branched-chain amino acid;
XX Corynebacterium; leucine; isoleucine; valine; medicine;
XX animal nutrition; ds.
XX Corynebacterium glutamicum.
XX Key Location/Qualifiers
XX CDS 101..856
XX FT /tag= a
XX FT /product= "brnF"
XX FT 853..1179
XX FT /tag= b
XX FT /product= "brnE"
XX EP1096010-A1.
XX PD 02-MAY-2001.
XX PF 11-OCT-2000; 2000BP-00122057.
XX PR 27-OCT-1999; 99DE-01051708.
XX PA (DEGS) DEGUSA AG.
XX PA (KERJ) FORSCHUNGSZENTRUM JUELICH GMBH.
XX PA Kennerknecht N, Eggeling L, Sahm H, Pfeifferle W;
XX PI

```

```

DR WPI; 2001-391595/42.
DR P-PSDB; AAB86247, AAB86248.
PT New export genes from corynebacterium bacteria, useful for increasing
PT fermentative production of branched-chain amino acids.
XX Claim 4 (1); Page 13; 23pp; German.
XX This invention describes a novel isolated polynucleotide (1) containing
XX at least one sequence that (i) is 70% identical with a sequence that
XX encodes at least one of 251 amino acid (aa) (3) or 108 aa (5)
XX polypeptides; (ii) encodes a polypeptide at least 70% identical with (3)
XX or (5); (iii) is the complement of (1) or (11), or (iv) contains at least
XX 15 consecutive bases from (1)-(11). The invention also describes (a)
XX protein derived from sequences of 753 bp (2) or 324 bp (4); (b)
XX corynebacterium microorganisms, especially Corynebacterium, transformed with
XX one or more (1), where these are replicative DNA; (c) production of
XX branched-chain L-aa by fermentation of corynebacterium bacteria in which the
XX brnE and/or brnF genes (or equivalent sequences) are amplified,
XX especially overexpressed; and (d) method for isolating the brnE and/or
XX brnF genes. (1) is used for transformation of corynebacterium bacteria being
XX used for fermentative production of branched-chain amino acids;
XX specifically leucine, isoleucine and valine, which are useful in medicine
XX and animal nutrition. (1) can also be used as source of primers and
XX probes for isolation of related sequences. Transformation with (1)
XX increases yield of branched-chain amino acids. This sequence encodes the
XX Corynebacterium glutamicum ATCC 14752 brnF and brnE proteins described in
XX the method of the invention
SQ Sequence 1271 BP; 265 A; 328 C; 318 G; 360 T; 0 U; 0 Other;
QY 100.0%; Score 753; DB 4; Length 1271;
QY Best Local Similarity 100.0%; Pred. No. 2,5e-22; Indels 0; Gaps 0;
QY Matches 753; Conservative 0; Mismatches 0;
QY 1 GTGCAAAAACGCAAGATTCATTCAGCTGAGGTGCGCATCGAGGAGCGCTG 60
DB 101 GTGCAAAAACGCAAGATTCATTCAGCTGAGGTGCGCATCGAGGAGCGCTG 160
QY 61 GAAACGATGATTAAGGTTATGCGCGCTACGAAATGCGCAAGGTAAACCTCCCT 120
DB 161 GAAACGATGATTAAGGTTATGCGCGCTACGAAATGCGCAAGGTAAACCTCCCT 220
QY 121 GCTGCAAGTTGGGCAATGATCCGATTTGATGCTGCTGTTATTCATAC 180
DB 221 GCTGCAAGTTGGGCAATGATCCGATTTGATGCTGCTGTTATTCATAC 280
QY 181 GGCTACGATGTTGGGCAAGCTGCTGTTTCCGCTGATTTTCCGGGCTCCACCGAA 240
DB 281 GGCTACGATGTTGGGCAAGCTGCTGTTTCCGCTGATTTTCCGGGCTCCACCGAA 340
QY 241 ATGCTGATTCAGCGCGCTGTTGTTGGGCGAGCGCCCTGGGCGCATCGCGCTAC 300
DB 341 ATGCTGATTCAGCGCGCTGTTGTTGGGCGAGCGCCCTGGGCGCATCGCGCTAC 400
QY 301 TTGCTGATGAATTCGCGCAGATATCTATGCGTTTTCATTCGCGTGCATGTGTCAAA 360
DB 401 TTGCTGATGAATTCGCGCAGATATCTATGCGTTTTCATTCGCGTGCATGTGTCAAA 460
QY 361 AACCCCATGCGCGTTTCTATTCGCGTTTTCGCGTTTTCGCGTTTTCGCGTTTTCGCG 420
DB 461 AACCCCATGCGCGTTTCTATTCGCGTTTTCGCGTTTTCGCGTTTTCGCGTTTTCGCG 520
QY 421 GCGGCGAGCGCGCGAGCGTGTGCGCGTGTGCGCGTGTGCGCGTGTGCGCGTGTGCGCG 480
DB 521 GCGGCGAGCGCGCGAGCGTGTGCGCGTGTGCGCGTGTGCGCGTGTGCGCGTGTGCGCG 580
QY 481 TCCTACTGCGTATTCGCGCGTGTGCGCGTGTGCGCGTGTGCGCGTGTGCGCGTGTGCG 540
DB 581 TCCTACTGCGTATTCGCGCGTGTGCGCGTGTGCGCGTGTGCGCGTGTGCGCGTGTGCG 640
QY 541 ATTAAGGCGCTCGAGTTCGCGCGTGTGCGCGTGTGCGCGTGTGCGCGTGTGCGCGTGTG 600

```


branched-chain L-a by fermentation of coryneform bacteria in which the
 CC brn and/or brn genes (or equivalent sequences) are amplified,
 CC especially overexpressed; and (d) method for isolating the brn and/or
 CC brn genes. (1) is used for transformation of coryneform bacteria being
 CC used for fermentative production of branched-chain amino acids,
 CC specifically leucine, isoleucine and valine, which are useful in medicine
 CC and animal nutrition. (1) can also be used as source of primers and
 CC probes for isolation of related sequences. Transformation with (1)
 CC increases yield of branched-chain amino acids. This sequence encodes the
 CC Corynebacterium glutamicum ATCC 14752 brnF protein described in the
 CC method of the invention

Sequence 753 BP; 140 A; 211 C; 194 G; 208 T; 0 U; 0 Other;

Query Match 100.0%; Score 753; DB 4; Length 753;

Best Local Similarity 100.0%; Pred. No. 2e-223; Mismatches 0; Indels 0; Gaps 0;

Matches 753; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 GTGCAAAAACGCAAGATTCATTAAGCTGAGGTGCGCATTCACAGCAGCCTG 60
 1 GTGCAAAAACGCAAGATTCATTAAGCTGAGGTGCGCATTCACAGCAGCCTG 60
 61 GAACGATGATTAAGGTTATCGGCGCTAGAAATCGCGCAAGGTCTAAAACCTCCTT 120
 61 GAACGATGATTAAGGTTATCGGCGCTAGAAATCGCGCAAGGTCTAAAACCTCCTT 120
 121 GGTGCAAGTTGGGCAATGACCGATGTTGCTTGGCTTGGCTTGAATAC 180
 121 GGTGCAAGTTGGGCAATGACCGATGTTGCTTGGCTTGGCTTGAATAC 180
 181 GGCTACGAATGAGGAGGAGCCCACTGTTTCCGGCTGATTTCCGGGCTCCACCGAA 240
 181 GGCTACGAATGAGGAGGAGCCCACTGTTTCCGGCTGATTTCCGGGCTCCACCGAA 240
 241 ATGCTGATCATCGCCTGCTTGGGCGCAGCGCCCTGGGCGCATTCGCGCTCACCA 300
 241 ATGCTGATCATCGCCTGCTTGGGCGCAGCGCCCTGGGCGCATTCGCGCTCACCA 300
 301 TTGCTGATGAACTTCCGCAAGTATTCATTCGCTTTCATTCGCGCTGATGTTGTTCAA 360
 301 TTGCTGATGAACTTCCGCAAGTATTCATTCGCTTTCATTCGCGCTGATGTTGTTCAA 360
 361 AACCCATTCGCGCTTTCATTCGCTTTCATTCGCTTTCATTCGCGCTGATGTTGTTCAA 420
 361 AACCCATTCGCGCTTTCATTCGCTTTCATTCGCTTTCATTCGCGCTGATGTTGTTCAA 420
 421 GCGGCGAGGCGCGAGGCTGGTGGCGGCTGAGCAATTCATGCAATGCAAAATAGGTTTAC 480
 421 GCGGCGAGGCGCGAGGCTGGTGGCGGCTGAGCAATTCATGCAATGCAAAATAGGTTTAC 480
 481 TCCTACTGAGTATTCGCGGCTTCACCGAGTGGCGATTCGCAATGCAATGCAATGCAATGCAAT 540
 481 TCCTACTGAGTATTCGCGGCTTCACCGAGTGGCGATTCGCAATGCAATGCAATGCAATGCAAT 540
 541 ATTAAGGCGCTGAGGTTGCGCTTTCATTCGCTTTCATTCGCTTTCATTCGCTTTCATTCGCTTTCAT 600
 541 ATTAAGGCGCTGAGGTTGCGCTTTCATTCGCTTTCATTCGCTTTCATTCGCTTTCATTCGCTTTCAT 600
 601 CGAAGCAAAAAGCAAGATTCCTCTCTGCTGCTGCAAGTTGAGCTTCAACCATTTGCTCTT 660
 601 CGAAGCAAAAAGCAAGATTCCTCTCTGCTGCTGCAAGTTGAGCTTCAACCATTTGCTCTT 660
 661 GTGCAATTCGAGGCAAGGCTTTCGCGGCGTGGCGATTCCTTGGGCTGTTGAC 720
 661 GTGCAATTCGAGGCAAGGCTTTCGCGGCGTGGCGATTCCTTGGGCTGTTGAC 720
 721 ATCCGTAATTCGAGGCAAGGCTTTCGCGGCGTGGCGATTCCTTGGGCTGTTGAC 720
 721 ATCCGTAATTCGAGGCAAGGCTTTCGCGGCGTGGCGATTCCTTGGGCTGTTGAC 720
 721 ATCCGTAATTCGAGGCAAGGCTTTCGCGGCGTGGCGATTCCTTGGGCTGTTGAC 720
 721 ATCCGTAATTCGAGGCAAGGCTTTCGCGGCGTGGCGATTCCTTGGGCTGTTGAC 720

RESULT 2
 AAH65254

AAH65254 standard; DNA; 753 BP.
 AAH65254;
 26-SEP-2001 (first entry)
 C glutamicum coding sequence fragment SEQ ID NO: 289.
 Corynebacterium glutamicum; amino acid synthesis; vitamin; saccharide;
 organic acid synthesis; ds.
 Corynebacterium glutamicum.
 EPI108790-A2.
 20-JUN-2001.
 18-DEC-2000; 2000EP-00127688.
 16-DEC-1999; 99JP-00377484.
 07-APR-2000; 2000JP-00159162.
 03-AUG-2000; 2000JP-00280988.
 (KJOW) KIOWA HAKKO KOGYO KK.
 Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;
 Tateishi N, Senoh A, Ikeda M, Ozaki A;
 WPI: 2001-376931/40.
 P-FSDB; AAG90035.
 Novel polynucleotides derived from Corynebacterium bacteria, for identifying
 mutation point of a gene, measuring expression of a gene, analyzing
 expression profile or pattern of a gene and identifying homologous gene.
 Claim 8; SEQ ID NO 289; 246bp + Sequence Listing; English.
 The present invention provides a number of nucleotide and protein
 sequences from the Corynebacterium glutamicum. These
 sequences are useful for identifying the mutation point of a gene derived from a
 mutant of Corynebacterium glutamicum, measuring expression amount and analyzing
 the expression profile or expression pattern of a gene derived from
 Corynebacterium glutamicum, and identifying a homologue of a gene derived from
 Corynebacterium glutamicum. Corynebacterium bacteria are useful for producing amino
 acids, nucleic acids, vitamins, saccharides and organic acids,
 particularly L-lysine. The present sequence is a nucleic acid described
 in the exemplification of the invention. Note: The sequence data for this
 patent did not form part of the printed specification, but was obtained
 in electronic format directly from the European Patent Office
 Sequence 753 BP; 140 A; 211 C; 194 G; 208 T; 0 U; 0 Other;
 Query Match 100.0%; Score 753; DB 5; Length 753;
 Best Local Similarity 100.0%; Pred. No. 2e-223; Mismatches 0; Indels 0; Gaps 0;
 Matches 753; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 1 GTGCAAAAACGCAAGATTCATTAAGCTGAGGTGCGCATTCACAGCAGCCTG 60
 1 GTGCAAAAACGCAAGATTCATTAAGCTGAGGTGCGCATTCACAGCAGCCTG 60
 61 GAACGATGATTAAGGTTATCGGCGCTAGAAATCGCGCAAGGTCTAAAACCTCCTT 120
 61 GAACGATGATTAAGGTTATCGGCGCTAGAAATCGCGCAAGGTCTAAAACCTCCTT 120
 121 GGTGCAAGTTGGGCAATGACCGATGTTGCTTGGCTTGGCTTGAATAC 180
 121 GGTGCAAGTTGGGCAATGACCGATGTTGCTTGGCTTGGCTTGAATAC 180
 181 GGCTACGAATGAGGAGGAGCCCACTGTTTCCGGCTGATTTCCGGGCTCCACCGAA 240
 181 GGCTACGAATGAGGAGGAGCCCACTGTTTCCGGCTGATTTCCGGGCTCCACCGAA 240
 241 ATGCTGATCATCGCCTGCTTGGGCGCAGCGCCCTGGGCGCATTCGCGCTCACCA 300

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 19, 2004, 10:56:53 ; Search time 350.669 Seconds
(without alignments)
9122.247 Million cell updates/sec

Title: us-10-608-504-6_COPY_101_853
Perfect score: 753
Sequence: 1 gtgcacaaacgcagagat.....tcttgggaagagctcctaaa 753

Scoring table: IDENTITY NUC
Gapop 10_0, Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues
Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_29yand4:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2003as:*
8: geneseqn2003bs:*
9: geneseqn2003cs:*
10: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	753	100.0	753	4 AAH21110	Aah21110 C. glutam
2	753	100.0	753	5 AAH65254	Aah65254 C. glutam
3	753	100.0	1271	4 AAH21109	Aah21109 C. glutam
4	753	100.0	1271	5 AAH21112	Aah21112 C. glutam
5	753	100.0	349860	4 AAH64966	Aah64966 C. glutam
6	553	73.4	564	7 ACA01969	ACA01969 C. glutam
7	273.6	36.3	397	4 AAH61693	Aah61693 C. glutam
8	273.6	36.3	778	4 AAH61695	Aah61695 C. glutam
9	127.4	16.9	711	5 AAH68373	Aah68373 C. glutam
10	127.4	16.9	711	5 AAH68373	Aah68373 C. glutam
11	127.4	16.9	309400	5 AAH68534	Aah68534 C. glutam
12	106	14.1	177	7 ACA01968	ACA01968 C. glutam
13	85	11.3	715	4 AAH61688	Aah61688 C. glutam
14	47	6.2	2000	7 ADA71938	Ada71938 Rice gene
15	41.2	5.5	10732	3 AAA10594	Aaa10594 Gene enco
16	40.8	5.4	7645	9 ADB59610	Ad59610 Human gen
17	40.2	5.3	4947	2 AAH68927	Aah68927 Teps17 po
18	39.8	5.0	768	9 ADC91639	Adc91639 E. faeciu
19	37.6	5.0	759	9 ACB61925	Ac61925 Photorhab
20	37.6	5.0	5059	2 AAH4332	Aah4332 Streptoc
21	37.6	5.0	110000	7 ACB7367_20	Ac7367_20 Continuat
22	37.6	5.0	243072	7 ACB5382	Ac5382 Photorhab
23	37.2	4.9	1024	6 ABX6366	Abx6366 Helicobac

24	36.8	4.9	1230	7 ACA42205	ACA42205 Prokaryot
25	36.4	4.8	751	2 AAV58489	AAV58489 3' fragme
26	36.4	4.8	751	2 AAV61146	AAV61146 3' CDNA s
27	36.4	4.8	751	3 AAA06252	Aaa06252 Human imm
28	36.4	4.8	751	3 AAB71157	Aab71157 Human pro
29	36.4	4.8	751	4 AAH93368	Aah93368 Human pro
30	36.4	4.8	751	4 AA633460	Aa633460 Human pro
31	36.4	4.8	751	4 AAH02433	Aah02433 Prostate
32	36.4	4.8	751	4 AAH84682	Aah84682 Human pro
33	36.4	4.8	751	5 ACA59269	ACA59269 Prostate
34	36.4	4.8	751	5 AAS10011	Aas10011 Human pro
35	36.4	4.8	751	6 ABL94832	Ab194832 Human N1-
36	36.4	4.8	751	6 ABB58541	Abb58541 Prostate
37	36.4	4.8	751	7 ACC94966	Acc94966 Prostate
38	36.4	4.8	751	9 ADB13462	Adb13462 Human pro
39	35.6	4.7	135638	7 ABX34289	Abx34289 S. atrool
40	35.4	4.7	1362	7 ADA70464	Ada70464 Rice gene
41	35.2	4.7	693	7 AB238247	Ab238247 N. gonorr
42	35.2	4.7	801	8 ADA31112	Ada31112 DNA encod
43	35.2	4.7	2001	4 AAH61084	Aah61084 P. putida
44	35.2	4.7	2028	4 AAH61083	Aah61083 P. putida
45	35.2	4.7	110000	4 AAH9682_33	Continuation (34 o

ALIGNMENTS

RESULT 1	AAH21110	AAH21110 standard; DNA; 753 BP.
AC	AAH21110;	
XX	05-SEP-2001 (first entry)	
DT		
XX		
DE	C. glutamicum brnF DNA.	
XX		
KW	L-amino acid production; brnF; brnG; branched-chain amino acid;	
KW	coryneform bacterium; leucine; isoleucine; valine; medicine;	
KW	animal nutrition; ds.	
OS	Corynebacterium glutamicum.	
XX		
PN	EP1096010-AI.	
XX		
XX	02-MAY-2001.	
PD		
XX		
PF	11-OCT-2000; 2000EP-00122057.	
XX		
FR	27-OCT-1999; 99DE-01051708.	
XX		
PA	(DEGS) DEGUSSA AG.	
XX	(KERJ) FORSCHUNGSZENTRUM JUELICH GMBH.	
PI	Kennernrecht N, Eggeling L, Sahm H, Pfeifferle W;	
DR	WPI; 2001-391595/42.	
XX	P-Psdb; AAB86247.	
PT	New export genes from coryneform bacteria, useful for increasing	
PT	fermentative production of branched-chain amino acids.	
XX		
PS	Claim 5; Page 14-15; 23pp; German.	
XX		
CC	This invention describes a novel isolated polynucleotide (I) containing	ACA42205 Prokaryot
CC	at least one sequence that (i) is 70% identical with a sequence that	AAV58489 3' fragme
CC	encodes at least one of 251 amino acid (aa) (3) or 108 aa (5)	AAV61146 3' CDNA s
CC	polypeptides; (ii) encodes a polypeptide at least 70% identical with (3)	Aaa06252 Human imm
CC	or (5); (iii) is the complement of (i) or (ii), or (iv) contains at least	Aab71157 Human pro
CC	15 consecutive bases from (i)-(iii). The invention also describes (a)	Aah93368 Human pro
CC	protein derived from sequences of 753 bp (2) or 324 bp (4); (b)	Aa633460 Human pro
CC	coryneform microorganisms, especially corynebacterium, transformed with	Aah02433 Prostate
CC	one or more (I), where these are replicative DNA; (c) production of	Aah84682 Human pro

Db 277489 GTGGTAATTCAGGTCAGGCTATTTCGGCGCTGATCTTCTGGTCTGTGACC 277548
QY 721 ATCCGTAATCTTCTTCTTGGAAAGGCTGCTAA 753
Db 277549 ATCCGTAATCTTCTTCTTGGAAAGGCTGCTAA 277581

RESULT 15

AX137083/c 397 bp DNA Linear PAT 30-MAY-2001
LOCUS AX137083
DEFINITION Sequence 7 from Patent EP1090993.
ACCESSION AX137083
VERSION AX137083.1 GI:14273428
KEYWORDS
SOURCE
ORGANISM Corynebacterium glutamicum
Corynebacterium glutamicum
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacteriaceae; Corynebacterium.

REFERENCE
1 Moeckel, B., Pfeifferle, W., Puenler, A., Kalinowski, J. and Bathe, B.
TITLE Nucleotide sequences coding for the lrp gene
JOURNAL Patent: EP 1090993-A 7 11-APR-2001;
Degussa AG (DE)

FEATURES
Source Location/Qualifiers
1..397
/organism="Corynebacterium glutamicum"
/mol_type="unassigned DNA"
/db_xref="taxon:1718"
/note="lrp-Teil 1"

ORIGIN

Query Match 36.3%; Score 273.6; DB 6; Length 397;
Best Local Similarity 98.6%; Pred. No. 3.8e-60;
Matches 276; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

1 GTGCAAAAAACGACAGAGATTTCATTCAGCTGAGGTGTGCGCATCCAGGACGACCTG 60
Db 284 GTGCAAAAAACGACAGAGATTTCATTCAGCTGAGGTGTGCGCATCCAGGACGACCTG 225
QY 61 GAACCGATGATTAAGTTATCGGCGCTACGAAATCGCGCAAGTCTAAACCTCCCTT 120
Db 224 GAACCGATGATTAAGTTATCGGCGCTACGAAATCGCGCAAGTCTAAACCTCCCTT 165
QY 121 GGTGCAAGTTGGGCAATGACCGATTGATTCGTTGCTCTCTTGGTATTCATAC 180
Db 164 GGTGCAAGTTGGGCAATGACCGATTGATTCGTTGCTCTCTTGGTATTCATAC 105
QY 181 GGTGCAAGTTGGGCAAGTCCCACTGTTTCCGGCTGATTTCCGGGCTCCACCGAA 240
Db 104 GGTGCAAGTTGGGCAAGTCCCACTGTTTCCGGCTGATTTCCGGGCTCCACCGAA 45
QY 241 ATGCTGTCATCGCCCTCGTTGTGGGGCAGCGCCCTG 280
Db 44 ATGCTGTCATCGCCCTCGTTGTGGGGCAGCGCCCTG 5

Search completed: April 19, 2004, 18:30:10
Job time : 3194.51 secs

```

/translation="MTSLRDLVSTNADAIADMSAFIEETVKGSGIGIALKGANG
AATKIDSDIVAKGSRLLPRLADSLDGMQYNGCTGATGLHENSSTALDAIIS
VADKNASINVEGLKATYKVGKRAKATIEDELPAIGLIERNKQ"
gene
    /gene="Cg10010"
    /complement(10104..11177)
    /gene="Cg10010"
    /codon_start=1
    /transl_table=11
    /product="Hypothetical protein"
    /protein_id="BAB97403.1"
    /db_xref="GI:21322774"
    /translation="MFMLAORTLPPIHTAPHLFVARVPHQIRATDADRTSLORDLELS
Query Match      100.0%; Score 753; DB 1; Length 340000;
Beet Local Similarity 100.0%; Pred. No. 1.2e-184;
Matches 753; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GTGCAAAAAACGCAAGATTCATTCAAGCCTGAGGTGTGCCATCCAGAGCCCTG 60
Db 276829 GTGCAAAAAACGCAAGATTCATTCAAGCCTGAGGTGTGCCATCCAGAGCCCTG 276888
QY 61 GAACCAAGATGATTAAGGTTATGCGGCTACGAAATCGCGCAAGGTCTAAAAACCTCCCTT 120
Db 276889 GAACCAAGATGATTAAGGTTATGCGGCTACGAAATCGCGCAAGGTCTAAAAACCTCCCTT 276948
QY 121 GCTGAGGTTTGGGCAATGACCCGATGCGTATGCGTGTGCTCTGCTTGATTCAATAC 180
Db 276949 GCTGAGGTTTGGGCAATGACCCGATGCGTATGCGTGTGCTCTGCTTGATTCAATAC 277008
QY 181 GGTACGAATGTTGGGCAAGCCCACTGTTTCCGGCTGATTTTGGCGGCTCCACGCA 240
Db 277009 GGTACGAATGTTGGGCAAGCCCACTGTTTCCGGCTGATTTTGGCGGCTCCACGCA 277068
QY 241 ATGCTGATCATGCGCCTCGTTGAGGCGCAGCGCCCTGAGCGCATGCGCTCACCA 300
Db 277069 ATGCTGATCATGCGCCTCGTTGAGGCGCAGCGCCCTGAGCGCATGCGCTCACCA 277128
QY 301 TTGCTGTGAATCTTCGCGCAAGTATTCTATGCGTTTCAATCCGCTGATGTGTCAA 360
Db 277129 TTGCTGTGAATCTTCGCGCAAGTATTCTATGCGTTTCAATCCGCTGATGTGTCAA 277188
QY 361 AACCCATTGCGCGTTTATTCGCTTTCGCGTTATGAGAGCAAGCCATGCACTACT 420
Db 277189 AACCCATTGCGCGTTTATTCGCTTTCGCGTTATGAGAGCAAGCCATGCACTACT 277248
QY 421 GCGGCAAGGCGCGAGGCTGTGCGGCGTGGCGATTAATCTCAATGCAATAGCGTTTAC 480
Db 277249 GCGGCAAGGCGCGAGGCTGTGCGGCGTGGCGATTAATCTCAATGCAATAGCGTTTAC 277308
QY 481 TCCTACTGGGATATTGCGGCGTCTCACCGGAGTGGGATGCGAGTGAATTCCTTTGAA 540
Db 277309 TCCTACTGGGATATTGCGGCGTCTCACCGGAGTGGGATGCGAGTGAATTCCTTTGAA 277368
QY 541 ATTAAGGCGCTCGAGTTTCGCGCTTTCCTCTCTTTGTCACGCTGACTTTGATTCCTGC 600
Db 277369 ATTAAGGCGCTCGAGTTTCGCGCTTTCCTCTCTTTGTCACGCTGACTTTGATTCCTGC 277428
QY 601 CGAAGCAAAAAGAGATTCCTCTCTCTGCTGCTCGAGGTTTGAAGTTACCATTCCTCT 660
Db 277429 CGAAGCAAAAAGAGATTCCTCTCTCTGCTGCTCGAGGTTTGAAGTTACCATTCCTCT 277488
QY 661 GTGGTAATTCAGATGAGCCCTATTTGCGGCGCTGATGCTCTTGCGTCTGTGACC 720
Db 277489 GTGGTAATTCAGATGAGCCCTATTTGCGGCGCTGATGCTCTTGCGTCTGTGACC 277548
QY 721 ATCCGCTACTTCTCTTGGGAAAAGGCTGCTAAA 753
Db 277549 ATCCGCTACTTCTCTTGGGAAAAGGCTGCTAAA 277581

```

```

LOCUS AX120085 349980 bp DNA linear PAT 11-MAY-2001
DEFINITION Sequence 1 from Patent EP1108730.
ACCESSION AX120085 AX114121
VERSION AX120085.1 GI:14036800
KEYWORDS
SOURCE
ORGANISM Corynebacterium glutamicum
            Corynebacterium glutamicum
            Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
            Corynebacterineae; Corynebacteriaceae; Corynebacterium.
REFERENCE
AUTHORS Nakagawa, S., Mizoguchi, H., Ando, S., Hayashi, M., Ochiai, K.,
        Yokoi, H., Tateishi, N., Senoh, A., Ikeda, M. and Ozaki, A.
TITLE Novel polynucleotides
JOURNAL Patent: EP 1108790-A 1 20-JUN-2001;
        KYOWA HAKKO KOGYO CO., LTD. (JP)
FEATURES
    source
        1..349980
        /organism="Corynebacterium glutamicum"
        /mol_type="genomic DNA"
        /db_xref="taxon:1718"
        /note="Seq 1 to long (3,309,400) split in 11, seq 1
        0,000,001 0,349,980 349,980"
ORIGIN
Query Match      100.0%; Score 753; DB 6; Length 349980;
Beet Local Similarity 100.0%; Pred. No. 1.2e-184;
Matches 753; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GTGCAAAAAACGCAAGATTCATTCAAGCCTGAGGTGTGCCATCCAGAGCCCTG 60
Db 276829 GTGCAAAAAACGCAAGATTCATTCAAGCCTGAGGTGTGCCATCCAGAGCCCTG 276888
QY 61 GAACCAAGATGATTAAGGTTATGCGGCTACGAAATCGCGCAAGGTCTAAAAACCTCCCTT 120
Db 276889 GAACCAAGATGATTAAGGTTATGCGGCTACGAAATCGCGCAAGGTCTAAAAACCTCCCTT 276948
QY 121 GCTGAGGTTTGGGCAATGACCCGATGCGTATGCGTGTGCTCTGCTTGATTCAATAC 180
Db 276949 GCTGAGGTTTGGGCAATGACCCGATGCGTATGCGTGTGCTCTGCTTGATTCAATAC 277008
QY 181 GGTACGAATGTTGGGCAAGCCCACTGTTTCCGGCTGATTTTGGCGGCTCCACGCA 240
Db 277009 GGTACGAATGTTGGGCAAGCCCACTGTTTCCGGCTGATTTTGGCGGCTCCACGCA 277068
QY 241 ATGCTGATCATGCGCCTCGTTGAGGCGCAGCGCCCTGAGCGCATGCGCTCACCA 300
Db 277069 ATGCTGATCATGCGCCTCGTTGAGGCGCAGCGCCCTGAGCGCATGCGCTCACCA 277128
QY 301 TTGCTGTGAATCTTCGCGCAAGTATTCTATGCGTTTCAATCCGCTGATGTGTCAA 360
Db 277129 TTGCTGTGAATCTTCGCGCAAGTATTCTATGCGTTTCAATCCGCTGATGTGTCAA 277188
QY 361 AACCCATTGCGCGTTTATTCGCTTTCGCGTTATGAGAGCAAGCCATGCACTACT 420
Db 277189 AACCCATTGCGCGTTTATTCGCTTTCGCGTTATGAGAGCAAGCCATGCACTACT 277248
QY 421 GCGGCAAGGCGCGAGGCTGTGCGGCGTGGCGATTAATCTCAATGCAATAGCGTTTAC 480
Db 277249 GCGGCAAGGCGCGAGGCTGTGCGGCGTGGCGATTAATCTCAATGCAATAGCGTTTAC 277308
QY 481 TCCTACTGGGATATTGCGGCGTCTCACCGGAGTGGGATGCGAGTGAATTCCTTTGAA 540
Db 277309 TCCTACTGGGATATTGCGGCGTCTCACCGGAGTGGGATGCGAGTGAATTCCTTTGAA 277368
QY 541 ATTAAGGCGCTCGAGTTTCGCGCTTTCCTCTCTTTGTCACGCTGACTTTGATTCCTGC 600
Db 277369 ATTAAGGCGCTCGAGTTTCGCGCTTTCCTCTCTTTGTCACGCTGACTTTGATTCCTGC 277428
QY 601 CGAAGCAAAAAGAGATTCCTCTCTCTGCTGCTCGAGGTTTGAAGTTACCATTCCTCT 660
Db 277429 CGAAGCAAAAAGAGATTCCTCTCTCTGCTGCTCGAGGTTTGAAGTTACCATTCCTCT 277488
QY 661 GTGGTAATTCAGATGAGCCCTATTTGCGGCGCTGATGCTCTTGCGTCTGTGACC 720

```


gene	/db_xref="taxon:1718"	
CDS	complement(385..711)	
	/gene="brnB"	
	complement(385..711)	
	/gene="brnB"	
	/note="LIV-E transporter component 1; branched-chain amino acid export protein"	
	/codon_start=1	
	/transl_table=11	
	/product="BrnB"	
	/protein_id="AA04685.1"	
	/db_xref="GI:21311380"	
	/translation="MTDPSCLLVAVCAVTPALRAVPELILKPLREDFVGMAMMPAGTILALITASTFRNAIDKTLTPELIVAVLITVAHLIGRRTLSVAGTIVFVGLVNL"	
gene	complement(708..1463)	
CDS	/gene="DrmF"	
	complement(708..1463)	
	/gene="brnF"	
	/note="LIV-E transporter component 2; branched-chain amino acid export protein"	
	/codon_start=1	
	/transl_table=11	
	/product="BrnF"	
	/protein_id="AA04686.1"	
	/db_xref="GI:21311381"	
	/translation="MKTQDEHSSLEVSPSKALPEDDKRYREIAGLTSIAGLGMPTIGIARGLVITGYEMWAPLPSGLIRAGSTEMLVIALVAGSARLISQIAHVNRFVAFASFPFLHVNKPIARFYSVALIDEAVAVAAPVAGSARLISQIAHSAVWFGGLVGAIAELIPEIKLEFALCSLVTLTLDSCRTKQIPELLAGSFTI ALVVIPOGALFPAALI FLGLITIRYFLGKAK"	
gene	1538..1993	
CDS	/gene="Lrp"	
	1538..1993	
	/gene="Lrp"	
	/note="Lrp"	
	/codon_start=1	
	/transl_table=11	
	/product="Lrp-like regulator"	
	/protein_id="AA04687.1"	
	/db_xref="GI:21311382"	
	/translation="MKDSDRAITAEISANRISNLAADKHLTPSPCRORLEAGTILGYSADIHPAVMNKGFEVTVLISNDRSTVNFSVAQHDVEVLEHLFGSPDYFRIGVADLEAYEQFLSHIQVGIKAKISRFAMKVKVPQV"	
ORIGIN		
Query Match	100.0%; Score 753; DB 1; Length 2105;	
Best Local Similarity	100.0%; Pred. No. 1e-184;	
Matches 753; Conservative	0; Mismatches 0; Indels 0; Gaps 0;	
QY	1 GTGCAAAAAACGACAGATTATTCAGCCTGAGAGTGTCCCATCCAGGACGCTTG	60
DB	1463 GTGCAAAAAACGACAGATTATTCAGCCTGAGAGTGTCCCATCCAGGACGCTTG	1404
QY	61 GAACCGATGATTAAGGTATGCGGCTACGAATGCGCAGAGTCTAAAACCTCCCT	120
DB	1403 GAACCGATGATTAAGGTATGCGGCTACGAATGCGCAGAGTCTAAAACCTCCCT	1344
QY	121 GTCGAGGTTTGGGCAATGACCGATTGGATTTGCTTTGCTCTTGGTTATTCATAC	180
DB	1343 GTCGAGGTTTGGGCAATGACCGATTGGATTTGCTTTGCTCTTGGTTATTCATAC	1284
QY	181 GGTACGAATGTGGGACGCCCACTGTTTCCGCGCTGATTTTGGCGGCTCCACCGAA	240
DB	1283 GGTACGAATGTGGGACGCCCACTGTTTCCGCGCTGATTTTGGCGGCTCCACCGAA	1224
QY	241 ATGCTGGTATCGCCTCGTGTGTGGGACGCGCCCTGCGGCGCATTCGCGTCAACCA	300
DB	1223 ATGCTGGTATCGCCTCGTGTGTGGGACGCGCCCTGCGGCGCATTCGCGTCAACCA	1164
QY	301 TTGCTGGTAACTTCGCCACGATTTCTATGCGTTTCAATCCCGCTGCATGTGTCAAA	360
DB	1163 TTGCTGGTAACTTCGCCACGATTTCTATGCGTTTCAATCCCGCTGCATGTGTCAAA	1104
QY	361 AACCCATTGCCCGTTTCTATTGCGTTTTCGCGCTTATCGACGAAGCTTACGAGTCACT	420
DB	1103 AACCCATTGCCCGTTTCTATTGCGTTTTCGCGCTTATCGACGAAGCTTACGAGTCACT	1044
QY	421 GCGGCGACGCGCGACGCTGTGCGGCGGCGACCTATGTCATGCAATGAAATAGGTTGAC	480
DB	1043 GCGGCGACGCGCGACGCTGTGCGGCGGCGACCTATGTCATGCAATGAAATAGGTTGAC	984
QY	481 TCCCTACTGGTATTCGCGCGCTGTCAACCGAGTGGCGCATGCAAGTGAATTCCTTTGAA	540
DB	983 TCCCTACTGGTATTCGCGCGCTGTCAACCGAGTGGCGCATGCAAGTGAATTCCTTTGAA	924
QY	541 ATTAAGGCGCTGAGTTCGCGCTTTCCTCTCTCTTGTGTCAAGCTGACTTGAATTCCTG	600
DB	923 ATTAAGGCGCTGAGTTCGCGCTTTCCTCTCTCTTGTGTCAAGCTGACTTGAATTCCTG	864
QY	601 CGAACGAAAAAGAGATCCCTTCTGTGCTGTGCGAGTTTGAGCTTCAACCATTCCTT	660
DB	863 CGAACGAAAAAGAGATCCCTTCTGTGCTGTGCGAGTTTGAGCTTCAACCATTCCTT	804
QY	661 GTGCTAATTCAGGTCAAGGCTTATTTGGGCGGCTGCTATCTTTGGGTCTGTTGAC	720
DB	803 GTGCTAATTCAGGTCAAGGCTTATTTGGGCGGCTGCTATCTTTGGGTCTGTTGAC	744
QY	721 ATCCGGTACTTCTTCTTGGGAAAGGCTGCTAA	753
DB	743 ATCCGGTACTTCTTCTTGGGAAAGGCTGCTAA	711
RESULT 13		
AP005274	340000 bp DNA linear BCT 08-AUG-2002	
LOCUS	Corynebacterium glutamicum ATCC 13032 DNA, complete genome, section	
DEFINITION	1/10.	
ACCESSION	AP005274 BA000036	
VERSION	AP005274.1 GI:21322764	
KEYWORDS		
SOURCE	Corynebacterium glutamicum ATCC 13032	
ORGANISM	Corynebacterium glutamicum ATCC 13032	
REFERENCE	Bacterial Actinobacteria; Actinobacteridae; Actinomycetales; Corynebacteriaceae; Corynebacterium.	
AUTHORS	Nakagawa, S.	
TITLE	Complete genomic sequence of Corynebacterium glutamicum ATCC 13032	
JOURNAL	Unpublished	
REFERENCE	2 (bases 1 to 340000)	
AUTHORS	Nakagawa, S.	
JOURNAL	Direct Submission	
COMMENT	Submitted (24-MAY-2002) Satoshi Nakagawa, Kyowa HAKKO Kogyo Co. Ltd., Tokyo Research Laboratories, 3-6-6, Asahi-machi, Machida, Tokyo 194-8533, Japan (E-mail: snakagawa@xanaden.com, Tel.:81-44-829-3031, Fax:81-44-813-1651)	
FEATURES	This sequence is conducted by collaboration of Kyowa HAKKO Kogyo Co. Ltd. And Kitasato University.	
source	Location/Qualifiers	
gene	1..340000	
	/organism="Corynebacterium glutamicum ATCC 13032"	
	/mol_type="genomic DNA"	
	/strain="ATCC 13032"	
	/db_xref="taxon:196627"	
	/note="ATCC 13032"	
CDS	1..1575	
	/gene="Cg10001"	
	1..1575	
	/gene="Cg10001"	
	/note="PF00308: Bacterial dnaA protein	
	TRG00362: dnaA: chromosomal replication initiator protein	
	dnaA"	
	/codon_start=1	
	/transl_table=11	
	/product="ATPase involved in DNA replication initiation"	
	/protein_id="BAB97394.1"	

Db 581 TCTACTGGGATTCGCGGCTCTACCGAGTGGCGATCGACAGATTGATTCCTTTGAA 640
Qy 541 ATTAAGGCGCTCGAGTTGCGCCCTTGTCTCTCTTTTGTACGCTGACTTTGGATTCTGCG 600
Db 641 ATTAAGGCGCTCGAGTTGCGCCCTTGTCTCTCTTTTGTACGCTGACTTTGGATTCTGCG 700
Qy 601 CGAAGCAAAAAGCAATCCCTTCTCTGCTGCTCGAGTTTGAGCTTACCATTTGCTCTT 660
Db 701 CGAAGCAAAAAGCAATCCCTTCTCTGCTGCTCGAGTTTGAGCTTACCATTTGCTCTT 760
Qy 661 GTGGTAATTCAGTCAGGCGCCCTATTGCGCGCTGATCTTCTTGGATCTGTGACC 720
Db 761 GTGGTAATTCAGTCAGGCGCCCTATTGCGCGCTGATCTTCTTGGATCTGTGACC 820
Qy 721 ATCCGTAATCTTCTTGTGGAAAAGCTGTAA 753
Db 821 ATCCGTAATCTTCTTGTGGAAAAGCTGTAA 853
RESULT 11
LOCUS BD014994 1271 bp DNA linear PAT 27-AUG-2002
DEFINITION Nucleotide sequence encoding the discharge of branched amino acid,
method of isolating the same and utilization thereof.
ACCESSION BD014994 GI:22555801
VERSION JP 2001169788-A/4.
KEYWORDS unclassified
SOURCE unclassified
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 1271)
AUTHORS Kernerhnecht,N., Sahm,H., Eggerling,L. and Pfeifferle,W.
TITLE Nucleotide sequence encoding the discharge of branched amino acid,
method of isolating the same and utilization thereof
JOURNAL Patent: JP 2001169788-A 4 26-JUN-2001;
DEGUSA HUELS AG,FORSCHUNGSZENTRUM JUELICH GMBH
COMMENT OS Corynebacterium glutamicum ATCC13032
PN JP 2001169788-A/4
PD 26-JUN-2001
PF 24-OCT-2000 JP 2000324315
PR 27-OCT-1999 DE 19951708.8
PI NICOLE KERNERHNECHT,HERMANN SAHM,LOTHAR EGGELING,WALTER P.
PFEIFFERLE
PC C12N15/09,C07K14/34,C12N1/21,C12P13/06,C12P13/06,C12P13/08//
PC C12N15/09,C12R1:15), (C12N1/21,C12R1:15), (C12P13/06,C12R1:15),
PC C12N15/00,
PC (C12N15/00,C12R1:15)
CC brnF
CC brnF
CC brnF
FH Key Location/Qualifiers
FT gene (101)..(853)
FT gene (853)..(1176).
FEATURES
source Location/Qualifiers
1..1271
/organism="unclassified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"
ORIGIN
Query Match 100.0%; Score 753; DB 6; Length 1271;
Best Local Similarity 100.0%; Pred. No. 1e-184;
Matches 753; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GTGCAAAAACGACAGATTCATTCAGGCTGAGGTGCGCATCCAGAGAGCCGCG 60
Db 101 GTGCAAAAACGACAGATTCATTCAGGCTGAGGTGCGCATCCAGAGAGCCGCG 160
Qy 61 GAACCAATGATTAAGGTTATCGCGCTGACGAATCGCGCAAGGTCTAAACCTCCCTT 120
Db 161 GAACCAATGATTAAGGTTATCGCGCTGACGAATCGCGCAAGGTCTAAACCTCCCTT 220
Qy 121 GCTGAGGTTTGGGCAATGACCGATTCGATTCGCTTGTGCTTGGTATTCAATAC 180

Db 221 GCTGAGGTTTGGGCAATGACCGATTCGATTCGCTTGTGCTTGGTATTCAATAC 280
Qy 181 GCTGAGGTTTGGGCAATGACCGATTCGATTCGCTTGTGCTTGGTATTCAATAC 240
Db 281 GCTGAGGTTTGGGCAATGACCGATTCGATTCGCTTGTGCTTGGTATTCAATAC 340
Qy 241 ATGCTGATGATCGCCCTCGTTGTGGGCGAGCGCCCTTGGGCGCATCGCGCTCAACA 300
Db 341 ATGCTGATGATCGCCCTCGTTGTGGGCGAGCGCCCTTGGGCGCATCGCGCTCAACA 400
Qy 301 TTGCTGATGATCGCCCTCGTTGTGGGCGAGCGCCCTTGGGCGCATCGCGCTCAACA 360
Db 401 TTGCTGATGATCGCCCTCGTTGTGGGCGAGCGCCCTTGGGCGCATCGCGCTCAACA 460
Qy 361 AACCCATGCGCGCTTCTATTGCTTGTGGGCGATTCGAGCAAGCTACAGCACTACT 420
Db 461 AACCCATGCGCGCTTCTATTGCTTGTGGGCGATTCGAGCAAGCTACAGCACTACT 520
Qy 421 GCGGCAAGCGCGCGCTGCTGCGGCGGCGCACTTATCTCAATGCAATAGCGTTTAC 480
Db 521 GCGGCAAGCGCGCGCTGCTGCGGCGGCGCACTTATCTCAATGCAATAGCGTTTAC 580
Qy 481 TCTACTGGGTTATTCGCGCGCTTCAACGAGTGGCGATGCGAGTTGATTCCTTTGAA 540
Db 581 TCTACTGGGTTATTCGCGCGCTTCAACGAGTGGCGATGCGAGTTGATTCCTTTGAA 640
Qy 541 ATTAAGGCGCTCGAGTTGCGCCCTTGTCTCTCTTGTGACGCTGACTTTGATTCTGCG 600
Db 641 ATTAAGGCGCTCGAGTTGCGCCCTTGTCTCTCTTGTGACGCTGACTTTGATTCTGCG 700
Qy 601 CGAAGCAAAAAGCAATCCCTTCTCTGCTGCTCGAGTTTGAGCTTACCATTTGCTCTT 660
Db 701 CGAAGCAAAAAGCAATCCCTTCTCTGCTGCTCGAGTTTGAGCTTACCATTTGCTCTT 760
Qy 661 GTGGTAATTCAGTCAGGCGCCCTATTGCGCGCTGATCTTCTTGGATCTGTGACC 720
Db 761 GTGGTAATTCAGTCAGGCGCCCTATTGCGCGCTGATCTTCTTGGATCTGTGACC 820
Qy 721 ATCCGTAATCTTCTTGTGGAAAAGCTGTAA 753
Db 821 ATCCGTAATCTTCTTGTGGAAAAGCTGTAA 853
RESULT 12
LOCUS AF454053/c 2105 bp DNA linear BCT 27-JUN-2002
DEFINITION Corynebacterium glutamicum BrnE (brnE), BrnF (brnF), and Lrp-like
regulator (Lrp) genes, complete cds.
ACCESSION AF454053
VERSION AF454053.1 GI:21311379
SOURCE Corynebacterium glutamicum
ORGANISM Corynebacterium glutamicum
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacterineae; Corynebacteriaceae; Corynebacterium.
REFERENCE 1 (bases 1 to 2105)
AUTHORS Kernerhnecht,N., Sahm,H., Yen,M.R., Patek,M., Saler Jr,M.H. Jr. and
Eggerling,L.
TITLE Export of L-isoleucine from Corynebacterium glutamicum: a
two-gene-encoded member of a new translocator family
JOURNAL J. Bacteriol. 184 (14), 3947-3956 (2002)
MEDLINE 22077265
PUBMED 12081967
REFERENCE 2 (bases 1 to 2105)
AUTHORS Kernerhnecht,N., Eggerling,L. and Sahm,H.
TITLE Direct Submission
JOURNAL Submitted (07-NOV-2001) IBT-1, Forschungszentrum, Leo-Brandt Str.,
Juelich 52425, Germany
FEATURES
source Location/Qualifiers
1..2105
/organism="Corynebacterium glutamicum"
/mol_type="genomic DNA"

CDS
853..1179
/note="unnamed protein product; brnE"
/codon_start=1
/transl_table=1
/protein_id="CAC39928.1"
/db_xref="GI:14273895"
/db_xref="EMBL:CAC39928"
/translation="MTDTSFSLVAVCAVTPALANPFLIKPERSQFGRKAM
MPAGILAILTASTFRSNAIDKTLTFGLIAVAITVAHLIGRRLLISVAGTIVF
GLVLPF"

Query Match 100.0%; Score 753; DB 6; Length 1271;
Best Local Similarity 100.0%; Pred. No. 1e-184;
Matches 753; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 GTGCAAAAACGACAGATTCATTCAAGCTGAGGAGTGTGCGCATCAAGAGCCCTG 60
101 GTGCAAAAACGACAGATTCATTCAAGCTGAGGAGTGTGCGCATCAAGAGCCCTG 160
61 GAACCAAGTATTAAGTTATCGGCGCTAGCAAAATCGCGCAAGTCTTAAACCTCCCT 120
161 GAACCAAGTATTAAGTTATCGGCGCTAGCAAAATCGCGCAAGTCTTAAACCTCCCT 220
121 GCTGAGGTTGGGATGTACCCGATTGATTCGTTGGTCTCTTGTATTCAATAC 180
221 GCTGAGGTTGGGATGTACCCGATTGATTCGTTGGTCTCTTGTATTCAATAC 280
181 GAGTACGAATGATGAGGAGCCCACTGTTTCCGAGCTGATTTCCGAGGCTCCACCGAA 240
281 GAGTACGAATGATGAGGAGCCCACTGTTTCCGAGCTGATTTCCGAGGCTCCACCGAA 340
241 ATGCTGATGATGAGGAGCTGATGAGGAGCCCACTGTTTCCGAGGCTCCACCGAA 300
341 ATGCTGATGATGAGGAGCTGATGAGGAGCCCACTGTTTCCGAGGCTCCACCGAA 400
301 TTGCTGATGATGAGGAGCTGATGAGGAGCCCACTGTTTCCGAGGCTCCACCGAA 360
401 TTGCTGATGATGAGGAGCTGATGAGGAGCCCACTGTTTCCGAGGCTCCACCGAA 460
361 AACCCATTGCGCGCTTCTATTCGCTTTCGCGCTTACGACGAAGCTTACGAGTCACT 420
461 AACCCATTGCGCGCTTCTATTCGCTTTCGCGCTTACGACGAAGCTTACGAGTCACT 520
421 GCGGCAAGGCGCGAGCTGATGAGGAGCCCACTGTTTCCGAGGCTCCACCGAA 480
521 GCGGCAAGGCGCGAGCTGATGAGGAGCCCACTGTTTCCGAGGCTCCACCGAA 580
481 TCTTACTGATGATGAGGAGCTGATGAGGAGCCCACTGTTTCCGAGGCTCCACCGAA 540
581 TCTTACTGATGATGAGGAGCTGATGAGGAGCCCACTGTTTCCGAGGCTCCACCGAA 640
641 ATTAAAGGCTGAGGAGCTGATGAGGAGCCCACTGTTTCCGAGGCTCCACCGAA 600
741 ATTAAAGGCTGAGGAGCTGATGAGGAGCCCACTGTTTCCGAGGCTCCACCGAA 700
601 CGAAGCAAAAACGACATCCCTTCTCTGCTGCTGCGAGGTTGAGCTTACCATTCCTCTT 660
701 CGAAGCAAAAACGACATCCCTTCTCTGCTGCTGCGAGGTTGAGCTTACCATTCCTCTT 760
661 GTGGAATTCAGGATCAGGCTGATTTGCGGCGCTGATGATTTGAGGCTGATGAGC 720
761 GTGGAATTCAGGATCAGGCTGATTTGCGGCGCTGATGATTTGAGGCTGATGAGC 820
721 ATCCGATCTTCTTCTGAGAAAGGCTGATAA 753
821 ATCCGATCTTCTTCTGAGAAAGGCTGATAA 853

RESULT 10
BD014991 1271 bp DNA linear PART 27-AUG-2002
LOCUS BD014991
DEFINITION Nucleotide sequence encoding the discharge of branched amino acid,

ACCESSION BD014991
VERSION BD014991.1 GI:22555798
KEYWORDS JP 2001169788-A/1.
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1 (bases 1 to 1271)
AUTHORS Kernerukunehito, N., Sahn, H., Eggering, L. and Pfeifferle, W.
TITLE Nucleotide sequence encoding the discharge of branched amino acid,
method of isolating the same and utilization thereof
PATENT: JP 2001169788-A 1 26-JUN-2001;
DEGUSA HUELS AG, FORSCHUNGSZENTRUM JUELICH GMBH
OS Corynebacterium glutamicum ATCC14752
PN JP 2001169788-A/1
PD 26-JUN-2001
PF 24-OCT-2000 JP 2000324315
PR 27-OCT-1999 DE 19951708.8
PI NICOLE KERNERUKUNEHITO, HERMANN SAHN, LOTHAR EGGERING, WALTER PI
PFEIFFERLE

FEATURES
source
1..1271
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"

ORIGIN

Query Match 100.0%; Score 753; DB 6; Length 1271;
Best Local Similarity 100.0%; Pred. No. 1e-184;
Matches 753; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 GTGCAAAAACGACAGATTCATTCAAGCTGAGGAGTGTGCGCATCAAGAGCCCTG 60
101 GTGCAAAAACGACAGATTCATTCAAGCTGAGGAGTGTGCGCATCAAGAGCCCTG 160
61 GAACCAAGTATTAAGTTATCGGCGCTAGCAAAATCGCGCAAGTCTTAAACCTCCCT 120
161 GAACCAAGTATTAAGTTATCGGCGCTAGCAAAATCGCGCAAGTCTTAAACCTCCCT 220
121 GCTGAGGTTGGGATGTACCCGATTGATTCGTTGGTCTCTTGTATTCAATAC 180
221 GCTGAGGTTGGGATGTACCCGATTGATTCGTTGGTCTCTTGTATTCAATAC 280
181 GAGTACGAATGATGAGGAGCCCACTGTTTCCGAGCTGATTTCCGAGGCTCCACCGAA 240
281 GAGTACGAATGATGAGGAGCCCACTGTTTCCGAGCTGATTTCCGAGGCTCCACCGAA 340
241 ATGCTGATGATGAGGAGCTGATGAGGAGCCCACTGTTTCCGAGGCTCCACCGAA 300
341 ATGCTGATGATGAGGAGCTGATGAGGAGCCCACTGTTTCCGAGGCTCCACCGAA 400
301 TTGCTGATGATGAGGAGCTGATGAGGAGCCCACTGTTTCCGAGGCTCCACCGAA 360
401 TTGCTGATGATGAGGAGCTGATGAGGAGCCCACTGTTTCCGAGGCTCCACCGAA 460
361 AACCCATTGCGCGCTTCTATTCGCTTTCGCGCTTACGACGAAGCTTACGAGTCACT 420
461 AACCCATTGCGCGCTTCTATTCGCTTTCGCGCTTACGACGAAGCTTACGAGTCACT 520
421 GCGGCAAGGCGCGAGCTGATGAGGAGCCCACTGTTTCCGAGGCTCCACCGAA 480
521 GCGGCAAGGCGCGAGCTGATGAGGAGCCCACTGTTTCCGAGGCTCCACCGAA 580
481 TCTTACTGATGATGAGGAGCTGATGAGGAGCCCACTGTTTCCGAGGCTCCACCGAA 540

Db 761 GTGTAATTCAGAGTCAGGCCCTATTGGCGCGCTGCTGATCTTTGGGCTGTGACC 820
QY 721 ATCCGTAATCTTCTTGGGAAAGGCTGCTAA 753
Db 821 ATCCGTAATCTTCTTGGGAAAGGCTGCTAA 853

RESULT 8
AX137709 1271 bp DNA linear PAT 30-MAY-2001
LOCUS Sequence 1 from Patent EP1096010.
AX137709
ACCESSION AX137709.1 GI:14273886
VERSION
KEYWORDS
SOURCE
ORGANISM Corynebacterium glutamicum
Corynebacterium glutamicum
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacteriaceae; Corynebacteriaceae; Corynebacterium.

REFERENCE
AUTHORS 1 Kernerhecht, N., Eggeling, L., Sahm, H. and Pfeifferle, W.
TITLE Nucleotide sequences coding for branched-chain amino acids export
proteins, method for isolating them and their use
JOURNAL Patent: EP 1096010-A 1 02-MAY-2001;
Degussa AG (DE); FORSCHUNGSZENTRUM JUELICH GMBH (DE)
FEATURES
source
1. 1271
/organism="Corynebacterium glutamicum"
/mol_type="unassigned DNA"
/db_xref="taxon:1718"
/note="ATCC14752"
101. 856
/note="unassigned protein product; brnf"
/codon_start=1
/transl_table=1
/protein_id="CAC39923.1"
/db_xref="GI:14273887"
/transl_table=1
/db_xref="REMBL:CA39923"
/translation="MKTQEHSLSEVSPKALEPDDKRYRREIAGLTSIAAGL
GMPIGIAFGLVIOGYEWMAAPLPSGLIFAGSTEMLVIALVGAAPLGAIALITLL
VNFHFVAFSPFLHVKNPIDARFYSVFLIDEAVALTAARPAWMSWRLISMOIAFH
SYVFGGLTGVAAIAELIPEIKGLEFALCSLVTLTDSCKTKQIPSLIAGLSFTI
ALVVIPOGALFAALIFLGLITIRYFLGRKAK"
853. 1179
/note="unassigned protein product; brnf"
/codon_start=1
/transl_table=1
/protein_id="CAC39924.1"
/db_xref="GI:14273888"
/db_xref="REMBL:CA39924"
/translation="MTDPSCLIVAVAVATFPAIRAVPFLIKPLAESOFVGMAM
WMPAGLIALITLSTFRSNATIDKTLTFLIIVAVITVAHLGSRITLISVAGTIVFA
GVNLP"

ORIGIN
Query Match 100.0%; Score 753; DB 6; Length 1271;
Best Local Similarity 100.0%; Pred. No. 1e-184;
Matches 753; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 GTGCAAAAAACGCAAGATTCATTCAAGCTTGAGGCTGCCCATCCAGGAGCCCTG 60
QY 101 GTGCAAAAAACGCAAGATTCATTCAAGCTTGAGGCTGCCCATCCAGGAGCCCTG 160
Db 61 GAACCAAGATTAAGGTTATCGCGCTAGCAAAATCGCGAAGGTCTAAAAACCTCCCT 120
QY 161 GAACCAAGATTAAGGTTATCGCGCTAGCAAAATCGCGAAGGTCTAAAAACCTCCCT 220
Db 121 GTGCAAGTTTGCGCATGATCCGATTTGATTTGCTTTGCTTTGTTATCAATAC 180
QY 221 GTGCAAGTTTGCGCATGATCCGATTTGATTTGCTTTGCTTTGTTATCAATAC 280
Db 181 GGCTACGAAGTGGGAGCGCCACATGTTTCCGGCCGATTTCCGGGGCTCCACCGAA 240
QY 281 GGCTACGAAGTGGGAGCGCCACATGTTTCCGGCCGATTTCCGGGGCTCCACCGAA 340

QY 241 ATGCTGTGATATGCGCCCTGCTGTGGGCGCAGGCCCGGAGGCGCATGCGGCTACGACA 300
Db 341 ATGCTGTGATATGCGCCCTGCTGTGGGCGCAGGCCCGGAGGCGCATGCGGCTACGACA 400
QY 301 TTGCTGTGAACCTTCGCGACAGTATTCATGCGTTTTCATTCGCCGTCATGTTGCTAA 360
Db 401 TTGCTGTGAACCTTCGCGACAGTATTCATGCGTTTTCATTCGCCGTCATGTTGCTAA 460
QY 361 AACCCATTCGCCGTTTTCATTCGCGTTTTCGCGTTTTCGCGTTTTCGCGTTTTCGCG 420
Db 461 AACCCATTCGCCGTTTTCATTCGCGTTTTCGCGTTTTCGCGTTTTCGCGTTTTCGCG 520
QY 421 GCGGCGAGCGCGCGAGGCTGTGCGGCGTGGCGCATTCATTCATGCAATAAGCGTTTAC 480
Db 521 GCGGCGAGCGCGCGAGGCTGTGCGGCGTGGCGCATTCATTCATGCAATAAGCGTTTAC 580
QY 481 TCCTACTGGGATTCGCGCGTCTCACCGGATGGCGATTCGCAAGTTGATTCCTTTTGA 540
Db 581 TCCTACTGGGATTCGCGCGTCTCACCGGATGGCGATTCGCAAGTTGATTCCTTTTGA 640
QY 541 ATTAAGGCGCTGAGTTGCGCCCTTTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 600
Db 641 ATTAAGGCGCTGAGTTGCGCCCTTTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 700
QY 601 CGAAGCAAAAAGCAGATCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 660
Db 701 CGAAGCAAAAAGCAGATCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 760
QY 661 GTGTAATTCAGATCAGGCGCCATTTGCGGCGCTGCTGATCTTTGAGTCTGTGACC 720
Db 761 GTGTAATTCAGATCAGGCGCCATTTGCGGCGCTGCTGATCTTTGAGTCTGTGACC 820
QY 721 ATCCGTAATCTTCTTGGGAAAGGCTGCTAA 753
Db 821 ATCCGTAATCTTCTTGGGAAAGGCTGCTAA 853

RESULT 9
AX137714 1271 bp DNA linear PAT 30-MAY-2001
LOCUS Sequence 6 from Patent EP1096010.
AX137714
ACCESSION AX137714.1 GI:14273893
VERSION
KEYWORDS
SOURCE
ORGANISM Corynebacterium glutamicum
Corynebacterium glutamicum
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacteriaceae; Corynebacteriaceae; Corynebacterium.

REFERENCE
AUTHORS 1 Kernerhecht, N., Eggeling, L., Sahm, H. and Pfeifferle, W.
TITLE Nucleotide sequences coding for branched-chain amino acids export
proteins, method for isolating them and their use
JOURNAL Patent: EP 1096010-A 6 02-MAY-2001;
Degussa AG (DE); FORSCHUNGSZENTRUM JUELICH GMBH (DE)
FEATURES
source
1. 1271
/organism="Corynebacterium glutamicum"
/mol_type="unassigned DNA"
/db_xref="taxon:1718"
/note="ATCC13032"
101. 856
/note="unassigned protein product; brnf"
/codon_start=1
/transl_table=1
/protein_id="CAC39927.1"
/db_xref="GI:14273894"
/db_xref="REMBL:CA39927"
/translation="MKTQEHSLSEVSPKALEPDDKRYRREIAGLTSIAAGL
GMPIGIAFGLVIOGYEWMAAPLPSGLIFAGSTEMLVIALVGAAPLGAIALITLL
VNFHFVAFSPFLHVKNPIDARFYSVFLIDEAVALTAARPAWMSWRLISMOIAFH
SYVFGGLTGVAAIAELIPEIKGLEFALCSLVTLTDSCKTKQIPSLIAGLSFTI
ALVVIPOGALFAALIFLGLITIRYFLGRKAK"

DEFINITION Sequence 1 from patent US 6613545.
ACCESSION AR391953
VERSION AR391953.1 GI:40115724
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE
TITLE 1 (bases 1 to 1271)
AUTHORS Kernerhnecht,N., Sahm,H., Eggeling,L. and Pfeifferle,W.
Nucleotide sequences coding for the export of branched chain amino acids, process for the isolation thereof and use thereof
JOURNAL Patent: US 6613545-A 1 02-SEP-2003;
FEATURES
source 1..1271
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN

Query Match 100.0%; Score 753; DB 6; Length 1271;
Best Local Similarity 100.0%; Pred. No. 1e-184;
Matches 753; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTGCAAAAACGAGAGATTCTTCAAGCTGAGAGTGTGCGCATCCAGGACCCCTG 60
DB 101 GTGCAAAAACGAGAGATTCTTCAAGCTGAGAGTGTGCGCATCCAGGACCCCTG 160
QY 61 GAACCATGATTAAGATTGCGGCTACGAATCGCGCAAGTCTAAAACTCCCTT 120
DB 161 GAACCATGATTAAGATTGCGGCTACGAATCGCGCAAGTCTAAAACTCCCTT 220
QY 121 GCTGCAAGTTTGGGCAATGACCCGATTGGATTCGGTTGCTCTTGGTTATTCATAC 180
DB 221 GCTGCAAGTTTGGGCAATGACCCGATTGGATTCGGTTGCTCTTGGTTATTCATAC 280
QY 181 GGCTACGAATGATGGGAGAGCCCACTGTTTCCGGCTGATTTTGGCGGGCTCCACGAA 240
DB 281 GGCTACGAATGATGGGAGAGCCCACTGTTTCCGGCTGATTTTGGCGGGCTCCACGAA 340
QY 241 ATGCTGATCATGCGCTCGTTGCGGAGAGCCCTGCGGCGCATCGCGCTCACCA 300
DB 341 ATGCTGATCATGCGCTCGTTGCGGAGAGCCCTGCGGCGCATCGCGCTCACCA 400
QY 301 TTGCTGATTAAGTTCCGCAAGATTCTATGCGTTTCAATCCGCTGATGCTGCA 360
DB 401 TTGCTGATTAAGTTCCGCAAGATTCTATGCGTTTCAATCCGCTGATGCTGCA 460
QY 361 AACCCATTGCCGTTTCTATTGCGTTTCCGCTTATCGAAGAGCTTACGATCACT 420
DB 461 AACCCATTGCCGTTTCTATTGCGTTTCCGCTTATCGAAGAGCTTACGATCACT 520
QY 421 GCGGCAAGGCGCGCAGGCTGCTGCGGCTGCGGCACTTATTCATATGCAATAGCGTTTAC 480
DB 521 GCGGCAAGGCGCGCAGGCTGCTGCGGCTGCGGCACTTATTCATATGCAATAGCGTTTAC 580
QY 481 TCCCTACTGAGTATTCGCGGCTGCTCAACGAGTGGCGATGCGAGATTGATTCCTTTGAA 540
DB 581 TCCCTACTGAGTATTCGCGGCTGCTCAACGAGTGGCGATGCGAGATTGATTCCTTTGAA 640
QY 541 ATTAAAGGCTCGAGTTGCGCTTGGCTCTCTTTGTCACGTCATCTTGGATTCCTGC 600
DB 641 ATTAAAGGCTCGAGTTGCGCTTGGCTCTCTTTGTCACGTCATCTTGGATTCCTGC 700
QY 601 CGAAGCAAAAAGAGATTCCTTCTGCTGCTGCGAGTTTGAAGTTACCATGCTCTT 660
DB 701 CGAAGCAAAAAGAGATTCCTTCTGCTGCTGCGAGTTTGAAGTTACCATGCTCTT 760
QY 661 GTGGTATTTCCAGTCAAGGCTCTTATTTGCGGCTGCTGATCTTCTTGGTCTGTGACC 720
DB 761 GTGGTATTTCCAGTCAAGGCTCTTATTTGCGGCTGCTGATCTTCTTGGTCTGTGACC 820
QY 721 ATCCGCTACTCTCTTGGGAAAGCTGTAA 753
DB 821 ATCCGCTACTCTCTTGGGAAAGCTGTAA 853

RESULT 7
AR391956 1271 bp DNA linear PAT 18-DEC-2003
LOCUS
DEFINITION Sequence 6 from patent US 6613545.
ACCESSION AR391956
VERSION AR391956.1 GI:40115727
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE
TITLE 1 (bases 1 to 1271)
AUTHORS Kernerhnecht,N., Sahm,H., Eggeling,L. and Pfeifferle,W.
Nucleotide sequences coding for the export of branched chain amino acids, process for the isolation thereof and use thereof
JOURNAL Patent: US 6613545-A 6 02-SEP-2003;
FEATURES
source 1..1271
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN

Query Match 100.0%; Score 753; DB 6; Length 1271;
Best Local Similarity 100.0%; Pred. No. 1e-184;
Matches 753; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTGCAAAAACGAGAGATTCTTCAAGCTGAGAGTGTGCGCATCCAGGACCCCTG 60
DB 101 GTGCAAAAACGAGAGATTCTTCAAGCTGAGAGTGTGCGCATCCAGGACCCCTG 160
QY 61 GAACCATGATTAAGATTGCGGCTACGAATCGCGCAAGTCTAAAACTCCCTT 120
DB 161 GAACCATGATTAAGATTGCGGCTACGAATCGCGCAAGTCTAAAACTCCCTT 220
QY 121 GCTGCAAGTTTGGGCAATGACCCGATTGGATTCGGTTGCTCTTGGTTATTCATAC 180
DB 221 GCTGCAAGTTTGGGCAATGACCCGATTGGATTCGGTTGCTCTTGGTTATTCATAC 280
QY 181 GGCTACGAATGATGGGAGAGCCCACTGTTTCCGGCTGATTTTGGCGGGCTCCACGAA 240
DB 281 GGCTACGAATGATGGGAGAGCCCACTGTTTCCGGCTGATTTTGGCGGGCTCCACGAA 340
QY 241 ATGCTGATCATGCGCTCGTTGCGGAGAGCCCTGCGGCGCATCGCGCTCACCA 300
DB 341 ATGCTGATCATGCGCTCGTTGCGGAGAGCCCTGCGGCGCATCGCGCTCACCA 400
QY 301 TTGCTGATTAAGTTCCGCAAGATTCTATGCGTTTCAATCCGCTGATGCTGCA 360
DB 401 TTGCTGATTAAGTTCCGCAAGATTCTATGCGTTTCAATCCGCTGATGCTGCA 460
QY 361 AACCCATTGCCGTTTCTATTGCGTTTCCGCTTATCGAAGAGCTTACGATCACT 420
DB 461 AACCCATTGCCGTTTCTATTGCGTTTCCGCTTATCGAAGAGCTTACGATCACT 520
QY 421 GCGGCAAGGCGCGCAGGCTGCTGCGGCTGCGGCACTTATTCATATGCAATAGCGTTTAC 480
DB 521 GCGGCAAGGCGCGCAGGCTGCTGCGGCTGCGGCACTTATTCATATGCAATAGCGTTTAC 580
QY 481 TCCCTACTGAGTATTCGCGGCTGCTCAACGAGTGGCGATGCGAGATTGATTCCTTTGAA 540
DB 581 TCCCTACTGAGTATTCGCGGCTGCTCAACGAGTGGCGATGCGAGATTGATTCCTTTGAA 640
QY 541 ATTAAAGGCTCGAGTTGCGCTTGGCTCTCTTTGTCACGTCATCTTGGATTCCTGC 600
DB 641 ATTAAAGGCTCGAGTTGCGCTTGGCTCTCTTTGTCACGTCATCTTGGATTCCTGC 700
QY 601 CGAAGCAAAAAGAGATTCCTTCTGCTGCTGCGAGTTTGAAGTTACCATGCTCTT 660
DB 701 CGAAGCAAAAAGAGATTCCTTCTGCTGCTGCGAGTTTGAAGTTACCATGCTCTT 760
QY 661 GTGGTATTTCCAGTCAAGGCTCTTATTTGCGGCTGCTGATCTTCTTGGTCTGTGACC 720

QY 181 GGCTACGAATGGTGGGACAGCCCACTGTTTCCGGCTGATTTTGGCGGGCTCCACCGAA 240
 DB 181 GGCTACGAATGGTGGGACAGCCCACTGTTTCCGGCTGATTTTGGCGGGCTCCACCGAA 240
 QY 241 ATGCTGATCATGCGCTCGTGTGGGCGAGCGCCCTGGGCGCATGCGCTCACCA 300
 DB 241 ATGCTGATCATGCGCTCGTGTGGGCGAGCGCCCTGGGCGCATGCGCTCACCA 300
 QY 301 TTGCTGTGAATCTTCCGCCACGATTTCTATGCGTTTCACTTCCCGCTGATGTGTCAA 360
 DB 301 TTGCTGTGAATCTTCCGCCACGATTTCTATGCGTTTCACTTCCCGCTGATGTGTCAA 360
 QY 361 AACCCATGCGCGCTTCTATTCGCTTTTCCGCTTATGACGAAAGCTTACAGTCACT 420
 DB 361 AACCCATGCGCGCTTCTATTCGCTTTTCCGCTTATGACGAAAGCTTACAGTCACT 420
 QY 421 GCGGCGAGCCCGCAGCGCTGTGCGGCGGCACTTATCTCATGCAATAGCGTTTAC 480
 DB 421 GCGGCGAGCCCGCAGCGCTGTGCGGCGGCACTTATCTCATGCAATAGCGTTTAC 480
 QY 481 TCCCTACTGGGATATTGGCGGCTTCACCGGAGTGGGAGATGCGAGATGATTCCTTTGA 540
 DB 481 TCCCTACTGGGATATTGGCGGCTTCACCGGAGTGGGAGATGCGAGATGATTCCTTTGA 540
 QY 541 ATTAAAGGCGCTCGAGTTGCGCCCTTGTCTCTCTTTGTCACGCTGACTTGGATTCTGC 600
 DB 541 ATTAAAGGCGCTCGAGTTGCGCCCTTGTCTCTCTTTGTCACGCTGACTTGGATTCTGC 600
 QY 601 CGAACGAAAGAGATTCCTTCTCTGCTGCTGCGAGGTTGAGCTTCCACATTCCTCT 660
 DB 601 CGAACGAAAGAGATTCCTTCTCTGCTGCTGCGAGGTTGAGCTTCCACATTCCTCT 660
 QY 661 GTGGTATTCAGATGAGGCGCTTATTTGGCGGCTGCTGATCTTCTTGGGCTGTGTGACC 720
 DB 661 GTGGTATTCAGATGAGGCGCTTATTTGGCGGCTGCTGATCTTCTTGGGCTGTGTGACC 720
 QY 721 ATCCGGTACTTCTTCTTGGGAAAGGCTGCTAAA 753
 DB 721 ATCCGGTACTTCTTCTTGGGAAAGGCTGCTAAA 753
 RESULT 5
 BD162490 753 bp DNA linear PAT 17-JAN-2003
 LOCUS BD162490 Novel polynucleotide.
 DEFINITION BD162490
 ACCESSION BD162490.1 GI:27868248
 VERSION JP 2002191370-A/289.
 KEYWORDS unclassified
 SOURCE unclassified
 ORGANISM unclassified
 1 (bases 1 to 753)
 Nakaawata,S., Mizoguchi,H., Ando,S., Hayaashi,M., Ochiai,K.,
 Yokoi,H., Tateishi,N., Senoo,A., Ikeda,M. and Ozaki,A.
 Novel polynucleotide
 Patent: JP 2002191370-A 289 09-JUL-2002;
 KYOWA HAKKO KOGYO CO LTD
 OS Corynebacterium glutamicum
 PN JP 2002191370-A/289
 PD 09-JUL-2002
 PF 15-DEC-2000 JP 2000405096
 PI SATOSHI NAKAGAWA,HIROSHI MIZOGUCHI,SEIKO ANDO,MIKIO HAYASHI,
 KEIKO OCHIAI,
 PI HARUHIKO YOKOI,NAOKO TATEISHI,AKIHIRO SENOO,MASATO IKEDA,AKIO
 OZAKI
 PC C12N15/09,C12N15/09,C07K14/34,C07K16/12,C07K16/40,C12M1/00,PC
 C12N1/15,
 PC C12N1/19,C12N1/21,C12N5/10,C12N9/00,C12N9/02,C12P1/40,C12P13/
 PC 04,C12P13/08,
 PC C12P19/00,C12P19/34,C12P21/02,C12Q1/37,C12Q1/68,G01N33/53,PC
 G01N33/56,
 PC G01N33/569,G01N33/68,G01N37/00/C12P21/08,C12N1/21,C12R1/15',
 PC (C12N1/21,C12R1/13),(C12N1/21,C12R1/01),(C12P13/08,C12R1/15),

PC C12N15/00.
 PC C12N5/00,C12N15/00
 CC Novel polynucleotide
 FH Key
 FT source
 FT Location/Qualifiers
 FEATURES
 source
 1..753
 /organism="unclassified"
 /mol_type="genomic DNA"
 /db_xref="taxon:32644"
 ORIGIN
 Query Match 100.0%; Score 753; DB 6; Length 753;
 Best Local Similarity 100.0%; Pred. No. 9.8e-185;
 Matches 753; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GTGCAAAAACGCAAGATTCATTCAGGCTGAGGTGTGCGCATTCAGAGCCCTG 60
 DB 1 GTGCAAAAACGCAAGATTCATTCAGGCTGAGGTGTGCGCATTCAGAGCCCTG 60
 QY 61 GAACCAATGATTAAGTTATCGGCGCTACGAATCGCGAAGTCTAAAACCTCCCT 120
 DB 61 GAACCAATGATTAAGTTATCGGCGCTACGAATCGCGAAGTCTAAAACCTCCCT 120
 QY 121 GCTGCAAGTTTGGGCGATGACCGGATTTGATTCGTTTGTCTTGTGTTAATAC 180
 DB 121 GCTGCAAGTTTGGGCGATGACCGGATTTGATTCGTTTGTCTTGTGTTAATAC 180
 QY 181 GGCTACGAATGGTGGGAGGCCACGTTTTCGGGCTGATTTTGGGCGCTCCACCGAA 240
 DB 181 GGCTACGAATGGTGGGAGGCCACGTTTTCGGGCTGATTTTGGGCGCTCCACCGAA 240
 QY 241 ATGCTGTGATTCGCGCTCGTGTGGGCGAGCGCCCTGGGCGCATTCGCGTACACCA 300
 DB 241 ATGCTGTGATTCGCGCTCGTGTGGGCGAGCGCCCTGGGCGCATTCGCGTACACCA 300
 QY 301 TTGCTGTGAATCTTCCGCCACGATTTCTATGCGTTTCACTTCCCGCTGATGTGTCAA 360
 DB 301 TTGCTGTGAATCTTCCGCCACGATTTCTATGCGTTTCACTTCCCGCTGATGTGTCAA 360
 QY 361 AACCCATGCGCGCTTCTATTCGCTTTTCCGCTTATGACGAAAGCTTACAGTCACT 420
 DB 361 AACCCATGCGCGCTTCTATTCGCTTTTCCGCTTATGACGAAAGCTTACAGTCACT 420
 QY 421 GCGGCGAGCCCGCAGCGCTGTGCGGCGGCACTTATCTCATGCAATAGCGTTTAC 480
 DB 421 GCGGCGAGCCCGCAGCGCTGTGCGGCGGCACTTATCTCATGCAATAGCGTTTAC 480
 QY 481 TCCCTACTGGGATATTGGCGGCTTCACCGGAGTGGGAGATGCGAGATGATTCCTTTGA 540
 DB 481 TCCCTACTGGGATATTGGCGGCTTCACCGGAGTGGGAGATGCGAGATGATTCCTTTGA 540
 QY 541 ATTAAAGGCGCTCGAGTTGCGCCCTTGTCTCTCTTTGTCACGCTGACTTGGATTCTGC 600
 DB 541 ATTAAAGGCGCTCGAGTTGCGCCCTTGTCTCTCTTTGTCACGCTGACTTGGATTCTGC 600
 QY 601 CGAACGAAAGAGATTCCTTCTCTGCTGCTGCGAGGTTGAGCTTCCACATTCCTCT 660
 DB 601 CGAACGAAAGAGATTCCTTCTCTGCTGCTGCGAGGTTGAGCTTCCACATTCCTCT 660
 QY 661 GTGGTATTCAGATGAGGCGCTTATTTGGCGGCTGCTGATCTTCTTGGGCTGTGTGACC 720
 DB 661 GTGGTATTCAGATGAGGCGCTTATTTGGCGGCTGCTGATCTTCTTGGGCTGTGTGACC 720
 QY 721 ATCCGGTACTTCTTCTTGGGAAAGGCTGCTAAA 753
 DB 721 ATCCGGTACTTCTTCTTGGGAAAGGCTGCTAAA 753
 RESULT 6
 AR391953 1271 bp DNA linear PAT 18-DEC-2003
 LOCUS AR391953

LOCUS AX137710 753 bp DNA linear PAT 30-MAY-2001
DEFINITION Sequence 2 from Patent EP1096010.
ACCESSION AX137710
VERSION AX137710.1 GI:14273889
KEYWORDS
SOURCE Corynebacterium glutamicum
ORGANISM Corynebacterium glutamicum
Bacteria; Actinobacteria; Actinomycetales;
Corynebacteriaceae; Corynebacteriales;
Corynebacterium.
REFERENCE
AUTHORS Kernerkecht, N., Eggeling, L., Sahm, H. and Pfeifferle, W.
TITLE Nucleotide sequences coding for branched-chain amino acids export
proteins, method for isolating them and their use
JOURNAL Patent: EP 1096010-A 2 02-MAY-2001;
Degussa AG (DE); FORSCHUNGSZENTRUM JUELICH GMBH (DE)
FEATURES
source location/Qualifiers
1..753
/organism="Corynebacterium glutamicum"
/mol_type="unassigned DNA"
/db_xref="taxon:1718"
/note="ATCC14752"
1..>753
/note="unassigned protein product; brnF"
/codon_start=1
/transl_table=1
/protein_id="CAC39925.1"
/db_xref="GI:14273890"
/db_xref="EMBL:CA39925"
/translation="MOKQEHHSLSLEVSPSRAALSPDDKRYRYEIAQGLTSLAQL
GMPYRIGALGLIVIOGYEMWAPYPSGLIRAGSLREMYVILVGAALGAIATLL
VNFHVFYAFSPFLVVKPILNFKYSVRLIDBAVYTAAPASMSRLISQIATH
SYWFGGLGVAIAELIPELKGLEFALCSFVTLTLDSCRTKQIPSLHAGSFTI
ALVVPGLAFNALIFGLITIRYFPLGKAK"
ORIGIN
Query Match 100.0%; Score 753; DB 6; Length 753;
Best Local Similarity 100.0%; Pred. No. 9.8e-185;
Matches 753; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GTGCAAAAAACGACAGATTCATTCAGCCTGAGGTGTCGCATCCAGGACGCTG 60
DB 1 GTGCAAAAAACGACAGATTCATTCAGCCTGAGGTGTCGCATCCAGGACGCTG 60
QY 61 GAACCGATGATTAAGATTATCGGCGCTACGAATTCGCGCATTAACCTCCCTT 120
DB 61 GAACCGATGATTAAGATTATCGGCGCTACGAATTCGCGCATTAACCTCCCTT 120
QY 121 GGTGCAAGTTGGGCAATGATACCGATGATGATGCTGCTCTGTTATTCATATC 180
DB 121 GGTGCAAGTTGGGCAATGATACCGATGATGATGCTGCTCTGTTATTCATATC 180
QY 181 GGCTACGAATGGTGGCAGCCCACTGTTTCCGCGCTATTTGCGGCGCTCCACGAA 240
DB 181 GGCTACGAATGGTGGCAGCCCACTGTTTCCGCGCTATTTGCGGCGCTCCACGAA 240
QY 241 ATGCTGATGATGCGCTGCTGTTGGGCGAGCGCCCTGAGGCGCATCGCGTCAACA 300
DB 241 ATGCTGATGATGCGCTGCTGTTGGGCGAGCGCCCTGAGGCGCATCGCGTCAACA 300
QY 301 TTGCTGATGATGCGCTGCTGTTGGGCGAGCGCCCTGAGGCGCATCGCGTCAACA 360
DB 301 TTGCTGATGATGCGCTGCTGTTGGGCGAGCGCCCTGAGGCGCATCGCGTCAACA 360
QY 361 AACCCATGCGCGCTGCTGTTGGGCGAGCGCTGATGAGAGAGCTACGAGTCACT 420
DB 361 AACCCATGCGCGCTGCTGTTGGGCGAGCGCTGATGAGAGAGCTACGAGTCACT 420
QY 421 GCGGCGAGCGCGCGAGGTGCTGCGCGTGGCGCATTAATCTCAATGCAATAGGTTTCA 480
DB 421 GCGGCGAGCGCGCGAGGTGCTGCGCGTGGCGCATTAATCTCAATGCAATAGGTTTCA 480
QY 481 TCTTACTGGGTATTCGCGCGGTCTCAGCGAGTGGCGCATGCAAGTTGATCCTTTGAA 540
DB 481 TCTTACTGGGTATTCGCGCGGTCTCAGCGAGTGGCGCATGCAAGTTGATCCTTTGAA 540

DB 481 TCTTACTGGGTATTCGCGCGGTCTCAGCGAGTGGCGCATGCAAGTTGATCCTTTGAA 540
QY 541 ATTAAGGCGCTCGAGTTGCGCCCTTTCCTCTGTTGTGACGCTGATTCGATTCGCG 600
DB 541 ATTAAGGCGCTCGAGTTGCGCCCTTTCCTCTGTTGTGACGCTGATTCGATTCGCG 600
QY 601 CGAAGCAAAAAAGAGATCCCTTCTCTGCTGCTGCGAGTTTGAAGTTCAACATGCTCT 660
DB 601 CGAAGCAAAAAAGAGATCCCTTCTCTGCTGCTGCGAGTTTGAAGTTCAACATGCTCT 660
QY 661 GTGTAATTCAGAGTCAAGGCGCTTATTTGGGCGCTGCTATCTTCTGGGTCTGTAAC 720
DB 661 GTGTAATTCAGAGTCAAGGCGCTTATTTGGGCGCTGCTATCTTCTGGGTCTGTAAC 720
QY 721 ATCCGTAATTCCTTTGGAAGAGCTGCTAA 753
DB 721 ATCCGTAATTCCTTTGGAAGAGCTGCTAA 753
RESULT 4
BD014992 753 bp DNA linear PAT 27-AUG-2002
LOCUS BD014992
DEFINITION Nucleotide sequence encoding the discharge of branched amino acid,
method of isolating the same and utilization thereof.
ACCESSION BD014992
VERSION BD014992.1 GI:22555799
KEYWORDS JP 2001169788-A/2.
SOURCE unidentified
ORGANISM unidentified
REFERENCE
AUTHORS Kernerkecht, N., Sahm, H., Eggeling, L. and Pfeifferle, W.
TITLE Nucleotide sequence encoding the discharge of branched amino acid,
method of isolating the same and utilization thereof
JOURNAL Patent: JP 2001169788-A 2 26-JUN-2001;
Degussa HUELZ AG, FORSCHUNGSZENTRUM JUELICH GMBH
COMMENT
OS Corynebacterium glutamicum ATCC14752
PN JP 2001169788-A/2
PD 26-JUN-2001
PF 24-OCT-2000 JP 2000324315
PR 27-OCT-1999 DE 19951708.8
PI NICOLE KENNERKECHT, HERMANN SAHM, LOTHAR EGGELING, WALTER PI
PFEIFFERLE
PC C12N15/09, C07K14/34, C12N1/21, C12P13/06, C12P13/06, C12P13/08//
PC C12N15/09, C12R1.15), (C12N1/21, C12R1.15), (C12P13/06, C12R1.15),
PC C12N15/00, C12R1.15)
PC (C12N15/00, C12R1.15)
CC brnF
FH key
FT CDS location/Qualifiers
1..753
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"
ORIGIN
Query Match 100.0%; Score 753; DB 6; Length 753;
Best Local Similarity 100.0%; Pred. No. 9.8e-185;
Matches 753; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GTGCAAAAAACGACAGATTCATTCAGCCTGAGGTGTCGCATCCAGGACGCTG 60
DB 1 GTGCAAAAAACGACAGATTCATTCAGCCTGAGGTGTCGCATCCAGGACGCTG 60
QY 61 GAACCGATGATTAAGATTATCGGCGCTACGAATTCGCGCATTAACCTCCCTT 120
DB 61 GAACCGATGATTAAGATTATCGGCGCTACGAATTCGCGCATTAACCTCCCTT 120
QY 121 GGTGCAAGTTGGGCAATGATACCGATGATGATGCTGCTCTGTTATTCATATC 180
DB 121 GGTGCAAGTTGGGCAATGATACCGATGATGATGCTGCTCTGTTATTCATATC 180
QY 181 GGCTACGAATGGTGGCAGCCCACTGTTTCCGCGCTATTTGCGGCGCTCCACGAA 240
DB 181 GGCTACGAATGGTGGCAGCCCACTGTTTCCGCGCTATTTGCGGCGCTCCACGAA 240
QY 241 ATGCTGATGATGCGCTGCTGTTGGGCGAGCGCCCTGAGGCGCATCGCGTCAACA 300
DB 241 ATGCTGATGATGCGCTGCTGTTGGGCGAGCGCCCTGAGGCGCATCGCGTCAACA 300
QY 301 TTGCTGATGATGCGCTGCTGTTGGGCGAGCGCCCTGAGGCGCATCGCGTCAACA 360
DB 301 TTGCTGATGATGCGCTGCTGTTGGGCGAGCGCCCTGAGGCGCATCGCGTCAACA 360
QY 361 AACCCATGCGCGCTGCTGTTGGGCGAGCGCTGATGAGAGCTACGAGTCACT 420
DB 361 AACCCATGCGCGCTGCTGTTGGGCGAGCGCTGATGAGAGCTACGAGTCACT 420
QY 421 GCGGCGAGCGCGCGAGGTGCTGCGCGTGGCGCATTAATCTCAATGCAATAGGTTTCA 480
DB 421 GCGGCGAGCGCGCGAGGTGCTGCGCGTGGCGCATTAATCTCAATGCAATAGGTTTCA 480
QY 481 TCTTACTGGGTATTCGCGCGGTCTCAGCGAGTGGCGCATGCAAGTTGATCCTTTGAA 540
DB 481 TCTTACTGGGTATTCGCGCGGTCTCAGCGAGTGGCGCATGCAAGTTGATCCTTTGAA 540

FEATURES
source
Location/Qualifiers
1..753
/organism="unknown"
/mol_type="genomic DNA"

Query Match 100.0%; Score 753; DB 6; Length 753;
Best Local Similarity 100.0%; Pred. No. 9.8e-185;
Matches 753; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

ORIGIN

1 GTGCAAAAAACGCAAGATTCATTCAGCCCTGAGGTGTGCGCATCCAGGAGCCCTG 60
1 GTGCAAAAAACGCAAGATTCATTCAGCCCTGAGGTGTGCGCATCCAGGAGCCCTG 60
61 GAACCAAGATGATTAAGATTATGCGGCTACGAAATCGCGCAAGTCTAAACCTCCCTT 120
61 GAACCAAGATGATTAAGATTATGCGGCTACGAAATCGCGCAAGTCTAAACCTCCCTT 120
121 GTCGACAGTTTGGGCAATGATACCGGATTTGCTGCTTTGCTTTGTTATTCATATAC 180
121 GTCGACAGTTTGGGCAATGATACCGGATTTGCTGCTTTGCTTTGTTATTCATATAC 180
181 GGTACGATGATGTTGGGAGCCCACTGTTTTCGCGCTATTTTCGCGGCTCCACCGAA 240
181 GGTACGATGATGTTGGGAGCCCACTGTTTTCGCGCTATTTTCGCGGCTCCACCGAA 240
241 ATGCTGTGATGCGCCCTGCTGTTGTTGGGCGAGCGCCCTGCGGCGCATCGCGCTCACACA 300
241 ATGCTGTGATGCGCCCTGCTGTTGTTGGGCGAGCGCCCTGCGGCGCATCGCGCTCACACA 300
301 TTGCTGTGATGCTTCGCGCATGATTTCTATGCGTTTCAATCCCGCTGATGTGCTCAA 360
301 TTGCTGTGATGCTTCGCGCATGATTTCTATGCGTTTCAATCCCGCTGATGTGCTCAA 360
361 AACCCATTGCGCGTTTCTATTGGGTTTGGGCTTATGCAAGAACTACGCACTACT 420
361 AACCCATTGCGCGTTTCTATTGGGTTTGGGCTTATGCAAGAACTACGCACTACT 420
421 GCGGCGAGGCCCGAGGCTGATGCGGCTGCGGCTGCACTTATCAGATGAAATAGCGTTTCA 480
421 GCGGCGAGGCCCGAGGCTGATGCGGCTGCGGCTGCACTTATCAGATGAAATAGCGTTTCA 480
481 TCCTACTGGGATTTGCGCGCTCTCAACCGAGTGGCGCATGCGAGTTGATTCCTTTGAA 540
481 TCCTACTGGGATTTGCGCGCTCTCAACCGAGTGGCGCATGCGAGTTGATTCCTTTGAA 540
541 ATTAAAGGCGCTCGAGTTGCGCGCTTGTCTCTCTTTGTCAAGCTGACTTTGATTCCTG 600
541 ATTAAAGGCGCTCGAGTTGCGCGCTTGTCTCTCTTTGTCAAGCTGACTTTGATTCCTG 600
601 CGAAGCAAAAAACGATTCCTTCTCTGCTGCTGCAAGTTTGAAGCTTCAACATTCCTT 660
601 CGAAGCAAAAAACGATTCCTTCTCTGCTGCTGCAAGTTTGAAGCTTCAACATTCCTT 660
661 GTGGAATATTCAGGCTCAGGCGCTTATTTGCGGCGCTGCTGATCTTTGAGTCTGTGACC 720
661 GTGGAATATTCAGGCTCAGGCGCTTATTTGCGGCGCTGCTGATCTTTGAGTCTGTGACC 720
721 ATCCGATATCTTCTTCTTGGGAAAGGCTGCTAAA 753
721 ATCCGATATCTTCTTCTTGGGAAAGGCTGCTAAA 753

RESULT 2
AX120373 753 bp DNA linear PAT 11-MAY-2001
LOCUS AX120373
DEFINITION Sequence 289 from Patent EP1108790.
ACCESSION AX120373
VERSION AX120373.1 GI:14037088
KEYWORDS
SOURCE Corynebacterium glutamicum
ORGANISM Corynebacterium glutamicum
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;

REFERENCE
AUTHORS
TITLE
JOURNAL
Kyowa HAKKO KOGYO CO., LTD. (JP)
Patent: EP 1108790-A 289 20-JUN-2001;
Novel polynucleotides
Yokoi, H., Tateishi, N., Senoh, A., Ikeda, M. and Ozaki, A.
Nakagawa, S., Mizoguchi, H., Ando, S., Hayaishi, M., Ochiai, K.,

FEATURES
source
Location/Qualifiers
1..753
/organism="Corynebacterium glutamicum"
/mol_type="unassigned DNA"
/db_xref="taxon:1718"

ORIGIN

Query Match 100.0%; Score 753; DB 6; Length 753;
Best Local Similarity 100.0%; Pred. No. 9.8e-185;
Matches 753; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 GTGCAAAAAACGCAAGATTCATTCAGCCCTGAGGTGTGCGCATCCAGGAGCCCTG 60
1 GTGCAAAAAACGCAAGATTCATTCAGCCCTGAGGTGTGCGCATCCAGGAGCCCTG 60
61 GAACCAAGATGATTAAGATTATGCGGCTACGAAATCGCGCAAGTCTAAACCTCCCTT 120
61 GAACCAAGATGATTAAGATTATGCGGCTACGAAATCGCGCAAGTCTAAACCTCCCTT 120
121 GTCGACAGTTTGGGCAATGATACCGGATTTGCTGCTTTGCTTTGTTATTCATATAC 180
121 GTCGACAGTTTGGGCAATGATACCGGATTTGCTGCTTTGCTTTGTTATTCATATAC 180
181 GGTACGATGATGTTGGGAGCCCACTGTTTTCGCGCTATTTTCGCGGCTCCACCGAA 240
181 GGTACGATGATGTTGGGAGCCCACTGTTTTCGCGCTATTTTCGCGGCTCCACCGAA 240
241 ATGCTGTGATGCGCCCTGCTGTTGTTGGGCGAGCGCCCTGCGGCGCATCGCGCTCACACA 300
241 ATGCTGTGATGCGCCCTGCTGTTGTTGGGCGAGCGCCCTGCGGCGCATCGCGCTCACACA 300
301 TTGCTGTGATGCTTCGCGCATGATTTCTATGCGTTTCAATCCCGCTGATGTGCTCAA 360
301 TTGCTGTGATGCTTCGCGCATGATTTCTATGCGTTTCAATCCCGCTGATGTGCTCAA 360
361 AACCCATTGCGCGTTTCTATTGGGTTTGGGCTTATGCAAGAACTACGCACTACT 420
361 AACCCATTGCGCGTTTCTATTGGGTTTGGGCTTATGCAAGAACTACGCACTACT 420
421 GCGGCGAGGCCCGAGGCTGATGCGGCTGCGGCTGCACTTATCAGATGAAATAGCGTTTCA 480
421 GCGGCGAGGCCCGAGGCTGATGCGGCTGCGGCTGCACTTATCAGATGAAATAGCGTTTCA 480
481 TCCTACTGGGATTTGCGCGCTCTCAACCGAGTGGCGCATGCGAGTTGATTCCTTTGAA 540
481 TCCTACTGGGATTTGCGCGCTCTCAACCGAGTGGCGCATGCGAGTTGATTCCTTTGAA 540
541 ATTAAAGGCGCTCGAGTTGCGCGCTTGTCTCTCTTTGTCAAGCTGACTTTGATTCCTG 600
541 ATTAAAGGCGCTCGAGTTGCGCGCTTGTCTCTCTTTGTCAAGCTGACTTTGATTCCTG 600
601 CGAAGCAAAAAACGATTCCTTCTCTGCTGCTGCAAGTTTGAAGCTTCAACATTCCTT 660
601 CGAAGCAAAAAACGATTCCTTCTCTGCTGCTGCAAGTTTGAAGCTTCAACATTCCTT 660
661 GTGGAATATTCAGGCTCAGGCGCTTATTTGCGGCGCTGCTGATCTTTGAGTCTGTGACC 720
661 GTGGAATATTCAGGCTCAGGCGCTTATTTGCGGCGCTGCTGATCTTTGAGTCTGTGACC 720
721 ATCCGATATCTTCTTCTTGGGAAAGGCTGCTAAA 753
721 ATCCGATATCTTCTTCTTGGGAAAGGCTGCTAAA 753

RESULT 3
AX137710

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 19, 2004, 10:57:18 / Search time 3191.51 Seconds

(without alignments)
10226.291 Million cell updates/sec

Title: US-10-608-504-6_COPY_101_853

Perfect score: 753

Sequence: 1 ggcacaaacgcagagat.....cttcggagaaagcgcctaaa 753

Scoring table: IDENTITY_NTC

Gapop 10.0, Gapext 1.0

Searched: 3470272 seqs, 2167151695 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database:

1: gb_ba:*

2: gb_ba:*

3: gb_ba:*

4: gb_ba:*

5: gb_ba:*

6: gb_ba:*

7: gb_ba:*

8: gb_ba:*

9: gb_ba:*

10: gb_ba:*

11: gb_ba:*

12: gb_ba:*

13: gb_ba:*

14: gb_ba:*

15: gb_ba:*

16: gb_ba:*

17: gb_ba:*

18: gb_ba:*

19: gb_ba:*

20: gb_ba:*

21: gb_ba:*

22: gb_ba:*

23: gb_ba:*

24: gb_ba:*

25: gb_ba:*

26: gb_ba:*

27: gb_ba:*

28: gb_ba:*

29: gb_ba:*

30: gb_ba:*

31: gb_ba:*

32: gb_ba:*

33: gb_ba:*

34: gb_ba:*

35: gb_ba:*

36: gb_ba:*

37: gb_ba:*

38: gb_ba:*

39: gb_ba:*

40: gb_ba:*

41: gb_ba:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	753	100.0	753	AR391954	AR391954 Sequence
2	753	100.0	753	AX120373	AX120373 Sequence
3	753	100.0	753	AX137710	AX137710 Sequence
4	753	100.0	753	BD014992	BD014992 Nucleotide
5	753	100.0	753	BD162490	BD162490 Novel pol
6	753	100.0	1271	AR391953	AR391953 Sequence
7	753	100.0	1271	AR391956	AR391956 Sequence
8	753	100.0	1271	AX137709	AX137709 Sequence
9	753	100.0	1271	AX137714	AX137714 Sequence
10	753	100.0	1271	BD014991	BD014991 Nucleotide
11	753	100.0	1271	BD014994	BD014994 Nucleotide
12	753	100.0	2105	AR454053	AR454053 Coryneb
13	753	100.0	340000	AP005274	AP005274 Coryneb
14	753	100.0	349980	AX120085	AX120085 Sequence
15	273.6	36.3	397	AX137083	AX137083 Sequence
16	273.6	36.3	778	AX137085	AX137085 Sequence
17	149.6	19.9	87340	AP005224	AP005224 Coryneb
18	132.6	17.6	53793	AX248361	AX248361 Coryneb
19	127.4	16.9	711	AX123492	AX123492 Sequence
20	127.4	16.9	711	BD165609	BD165609 Novel pol
21	127.4	16.9	309400	AX127153	AX127153 Sequence
22	127.4	16.9	325651	AP005283	AP005283 Coryneb
23	85	11.3	715	AX137077	AX137077 Sequence
24	85	11.3	715	BD013995	BD013995 Novel nuc
25	59.8	7.9	349926	AX571660	AX571660 Wolinella
26	58.6	7.8	756	AR366287	AR366287 Sequence
27	57.2	7.6	9888	AE011050	AE011050 Methanosa
28	54.4	7.2	856	HSB325448	HSB325448 Homo sapi
29	51.8	6.9	11025	AE013585	AE013585 Methanosa
30	51	6.8	308015	AE016783	AE016783 Pseudomon
31	48	6.4	301838	AE017209	AE017209 Geobacter
32	47	6.2	2000	AX655393	AX655393 Sequence
33	46.6	6.2	3619	BCR410307	BCR410307 Erythrina c
34	46.6	6.2	10733	AE013695	AE013695 Yersinia
35	46.6	6.2	220050	AX141456	AX141456 Yersinia
36	46.2	6.1	77218	166494	166494 Sequence 14
37	46	6.1	215050	AL646084	AL646084 Ralstonia
38	45.8	6.1	20342	AE008888	AE008888 Salmonell
39	45.4	6.0	348624	AX640441	AX640441 Bordetell
40	45.4	6.0	348866	AX640426	AX640426 Bordetell
41	44.6	5.9	329709	AP002997	AP002997 Mesorhizo
42	44.2	5.9	208524	AE016925	AE016925 Chromobac
43	44.2	5.9	264050	AL6527279	AL6527279 Salmonell
44	44.2	5.9	300592	AE016846	AE016846 Salmonell
45	43.4	5.8	346294	AP002999	AP002999 Mesorhizo

ALIGNMENTS

RESULT 1	AR391954	AR391954	753 bp	DNA	linear	PAT 18-DEC-2003
LOCUS	AR391954	Sequence 2 from patent US 6613545.				
DEFINITION	AR391954	Sequence 2 from patent US 6613545.				
ACCESSION	AR391954	Sequence 2 from patent US 6613545.				
VERSION	AR391954.1	GI:40115725				
KEYWORDS	Unknown.					
SOURCE	Unknown.					
ORGANISM	Unknown.					
REFERENCE	1 (bases 1 to 753)					
AUTHORS	Kennethrecht,N., Salm,H., Eggeling,I., and Pfefferle,W.					
TITLE	Nucleotide sequences coding for the export of branched chain amino acids, process for the isolation thereof and use thereof					
JOURNAL	Patent: US 6613545-A 2 02-SEP-2003;					

Qy	119	ATTCATTCAGCGCTGGAGAGTGTGGCATTCGACGAGCGCTGGAACAGATGATTAAGGT	178
	825	MDMMMMMMMMNDKKKKNDKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKK	884
Qy	179	TATGGGCGTACGAATATGCGGACGTCTAAAACTCCCTTGCTGCAGGTGGGCGATG	238
Db	885	KKKKKKDAKK	944
Qy	239	TACCGATGTGATATGGCGTTGGATCTTGTCATACATGCGCTACGATGTGGGCA	298
Db	945	KKKKKKKKTKKKTTTTTTTHTTTTMMHHTTTTTTTTTTTMMTTTTTTHKKKKKKKK	1004
Qy	299	GCCCCACTGTTTTCCGGCCTGATTTTTCGGGGGCTCCACCGAAATGCTGGTCATGGCCCTC	358
Db	1005	KNDNNMMMMKK	1064
Qy	359	GTTGGGCGGAGCGCCCTCGGCGCGCATCGCGTCAACACATTCGTGGTAACTTCGC	418
Db	1065	KKKKKKKCCCBKBCGCCCKKXCCBKXCCYCKKKKKKKKKKKYCCCBCKKKCBGVC	1124
Qy	419	CACGATTCATGCGTTTCANT	441
Db	1125	CCCKKKKCBCKKKKKKKCCCKKK	1147

RESULT 15

CB657749	835 bp	mRNA	linear	EST 09-APR-2003
LOCUS	CB657749			
DEFINITION	OSJunc13F02.f OSJunc Oryza sativa (japonica cultivar-group) cDNA clone OSJunc13F02 5', mRNA sequence.			

ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM
CB657749	CB657749.1	GI:29661474	EST.	<i>Oryza sativa</i> (japonica cultivar-group)
				<i>Oryza sativa</i> (japonica cultivar-group)

REFERENCE
AUTHORS
TITLE
JOURNAL

1 (bases 1 to 835)
Jantasuyarat, C., Lu, G., Gowda, M., Hatfield, J., Zhou, B., Mazur, E.,
Kudrna, D., Dean, R., Soderlund, C., Wieg, R., and Wang, G.
Large-scale identification of ESTs involved in the interaction
between rice and Magnaporthe oryzae
Unpublished (2003)

JOURNAL Unpublished (2003)
COMMENT Contact: Rod Wing

Arizona Genomics Institute
University of Arizona
Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
85721-0088, USA

Tel: 520 626 3967
Fax: 520 621 9288

Email: <http://genome.arizona.edu>
PCR Primers

```
FORWARD: gta aaa cga cgg cca gtg
BACKWARD: gga aac agc tat gac cat g
```

Plate: 13 row: F column: 02
Seq primer: gta aaa cga cgg cca gtc.

FEATURES	Location/Qualifiers
source	1. .835

/organism="Oryza sativa (japonica cultivar-group)"

```

/mol_type="mRNA"
/cultivar="Niponbare"

```

```

/db_xref="taxon:39947"
/clone="OSJNEc13F02"

```

```
/tissue_type="Leaf"
/dev_stage="3 week"
```

```

/lab_host="DH10B"
/clone_lib="OSJNEC"

```

```

/note="Vector: pBluescript II KS +; Site_1: EcoRI; Site_2:
XhoI; 6 hrs after inoculation with Rice Blast (C9240-1)"

```

ORIGIN

Query Match	3.0%;	Score 38;	DB 14;	Length 835;
Best Local Similarity	51.1%;	Pred. No. 28;		
Matches	89;	Conservative	0;	Mismatches 85;
			Indels	0;
			Gaps	0;

QY 352 CGCCCTCGTTGTGGGCGCAGCGCCCTTGGGGGCCCATGCGGTACACCAATTGTGGTGA 411

Db 79 CGACGCGGCTCCGGTGGCCGGCTGCTCGGGGCTTTCGCGCGTGGCGAGATGCCCTC 138

QY 412 CTTTCGCCACGATATTCTATAGGTTTTCATTCCCGCTGCATGTGTGTAATAAACCCTATGC 471

Db 139 ATGCGCGCGGGGTTCGACCGGATCGCGCACCCCGACTCGGCTCGGACCGCGTGC 198

QY 472 CGATTTCATTGCTTTTCGCGCTTATCGACGAAGCTTACGCACTATCGGGC 525

Db 199 CTCGCGCTACGGGCGCTCTTCGCGCAAGGAGTGGCTGGCGCACCTCTGCGGC 252

Search completed: April 17, 2004, 20:01:48
Job time : 416.78 secs

FEATURES
source

http://image.llnl.gov
Plate: LLM9410 row: k column: 05
High quality sequence stop: 385.
Location/Qualifiers
1..859
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4149244"
/issue_type="G10blastoma with EGFR amplification"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NCI CGAP Brn4"
/note="Organ: Brain; Vector: pCMV-SPORT6; Site: 1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dr.
Average insert size 1.57 kb. Constructed by Life
Technologies. Note: this is a NCI CGAP Library."

ORIGIN

Query Match 3.1%; Score 39; DB 10; Length 859;
Best Local Similarity 51.4%; Pred. No. 15;
Matches 90; Conservative 0; Mismatches 85; Indels 0; Gaps 0;

QY 760 TGTGTAATTCAGGTCAGCCCTATTGCGGCGCTGATCTTGGGTCTGTAC 819
Db 792 TGTGTTCTTCTGGGTGTCCTTCTGTCGCGCGTCTGCTGTTTCTGCG 733
QY 820 CATCCGTAATCTTCTTGGGAAAGGCTGCTTAATGACACTGATTTCTCTC 879
Db 732 TCTGATATATTCCTGCTTCTGCTCGGTCCTTCTACTCGTTTGTGTTTGT 673
QY 880 CTGTGTCGAGTATGTCAGTCATTAATTTGCGCGCGGCGGTTCCGTTCT 934
Db 672 GTTGTGTCCTGATGTCGCTGCTTGTGTTATTCCTCCGTGCTCGTCT 618

RESULT 13
CNS0091P 925 bp DNA linear GSS 03-JUN-1999
LOCUS Drosophila melanogaster genome survey sequence Tera end of BAC #
DEFINITION BACR19D16 of RCI-98 library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
ACCESSION CNS0091P
VERSION AL053013.1 GI:4934461
KEYWORDS GSS.
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 925)
Genoscope.
Direct Submission
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : segref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see <http://www.fruitfly.org> The BDGP Drosophila
melanogaster BAC library was prepared by Karuoto Osoegawa and
Aaron Mammeter in Peter de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RCI-98 and was constructed by partial
EcoRI digestion of Drosophila DNA provided by the BDGP from the
isogenic strain Y2; cn bw sp, the same strain used for the BDGP's
p1 and EST libraries. A more detailed description of the library
and how to order individual BAC clones, the entire library, or
filters for hybridization from the BACPAC Resource Center can be
found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
Location/Qualifiers
1..925
/organism="Drosophila melanogaster"

FEATURES
source

http://image.llnl.gov
Plate: LLM9410 row: k column: 05
High quality sequence stop: 385.
Location/Qualifiers
1..859
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4149244"
/issue_type="G10blastoma with EGFR amplification"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NCI CGAP Brn4"
/note="Organ: Brain; Vector: pCMV-SPORT6; Site: 1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dr.
Average insert size 1.57 kb. Constructed by Life
Technologies. Note: this is a NCI CGAP Library."

FEATURES
source

http://image.llnl.gov
Plate: LLM9410 row: k column: 05
High quality sequence stop: 385.
Location/Qualifiers
1..859
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4149244"
/issue_type="G10blastoma with EGFR amplification"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NCI CGAP Brn4"
/note="Organ: Brain; Vector: pCMV-SPORT6; Site: 1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dr.
Average insert size 1.57 kb. Constructed by Life
Technologies. Note: this is a NCI CGAP Library."

ORIGIN

Query Match 3.1%; Score 39; DB 29; Length 925;
Best Local Similarity 14.2%; Pred. No. 15;
Matches 37; Conservative 117; Mismatches 106; Indels 0; Gaps 0;

QY 295 GGACAGCCCACTTTTCCGCTGATTTTCGGGCGCTCCACCAATGCTGATCC 354
Db 638 SSSSSSTSSSTSSSTSSSSSSSSSSSSSSSTSSSTSSSSSSSSSSSS 697
QY 355 CTTGTTGGGGCGGAGCGCCCTGGCGCCATCGGCTCACCACTGCTGACTT 414
Db 698 SSSSSSTSSSSSVSSSGKSTBSGSSSSSSSSSSSSSTSSSSSSSSST 757
QY 415 CCGCCAGATTTCTACGCTTTCATTCCTCCGTCGATGTCACAAACCCATTGCCG 474
Db 758 CCGCTCCSYSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSS 817
QY 475 TTCTATTCGATTTTCGCGCTTATTCAGACGCTTACGACGCTACGCGCCGCG 534
Db 818 YBMCTSTSGCGSSSSSGKGVTKCGCGCGSSSTNGMBGTSSACSSSSSSSVSS 877
QY 535 AGCGTGTCGCGCTGCGGAC 554
Db 878 SSKSSASSSSSVSSSGSSGS 897

RESULT 14
EX361080 1201 bp mRNA linear EST 05-MAY-2003
LOCUS EX361080 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
DEFINITION clone CSOD1079YN16 3-PRIME, mRNA sequence.
ACCESSION EX361080
VERSION EX361080.1 GI:30374504
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1201)
AUTHORS Li, W.B., Gruber, C., Jesse, J., and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: segref@genoscope.cns.fr; Web: www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. Contact: Feng Liang Email: fliang@lifetech.com URL :
<http://fulllength.invitrogen.com/invitrogen/1600>
Faraday Avenue, Genoscope, sequence ID : CSOD1079D608P1.
Location/Qualifiers
1..1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSOD1079YN16"
/issue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo (dr)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

FEATURES

source

ORIGIN

Query Match 3.0%; Score 38.6; DB 13; Length 1201;
Best Local Similarity 10.5%; Pred. No. 22;
Matches 34; Conservative 147; Mismatches 142; Indels 0; Gaps 0;

COMMENT

COMMENT

Contact : Genoscope
Genoscope - Centre National de Sequencage
BP 131 91006 EVRY cedex - France
Email : segred@genoscope.cns.fr Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. Contact : Feng Liang Email : liang@lifetech.com URL :
<http://fulllength.invitrogen.com/> INVITROGEN CORP 1600
Farday Avenue, Genoscope, sequence ID : C10B022A07P7P1

ORIGIN

Query Match	3.1%	Score 39.4	DB 13	Length 885
Similarity	1.5%	Pred. No. 11		
Best Local	7	Mismatches 147	Indels 0	Gaps 0
Matches				

RESULT 11	CNS005NG/c	LOCUS	CNS005NG	995 bp	DNA	linear	GSS 03-JUN-1995
DEFINITION	Drosophila melanogaster genome survey sequence 1513 end of BAC # BAC11264 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.						
ACCESSION	AL060428						
VERSION	AL060428.1						
KEYWORDS	GSS.						
SOURCE	Drosophila melanogaster (fruit fly)						
ORGANISM	Drosophila melanogaster						
	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;						

REFERENCE
AUTHORS
TITLE
JOURNAL

COMMENT

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
1	(bases 1 to 995)	Ephydriidae, Drosophilidae; Drosophila.		
	Genoscope.			
	Direct Submission			
	Submitted (02-JUN-1999)	Genoscope - Centre National de Sequencage		
	Bp 101 91006 EVRY cedex - FRANCE	(E-mail : seqref@genoscope.cns.fr		
	- Web : www.genoscope.cns.fr)			
	Determination of this BAC-end sequence was carried out as part of a			
	collaboration with the Berkeley Drosophila Genome Project (BDGP).			
	The BDGP is constructing a physical map of the Drosophila			
	melanogaster genome using these BACs. For further information			
	please see http://www.fruitfly.org The BDGP Drosophila			
	melanogaster BAC library was prepared by Kazuhiro Oosagawa and			
	Aaron Mammone in Pierer de Jong's laboratory in the Department of			
	Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,			
	NY. The library is named RPCC-98 and was constructed by partial			
	EcoRI digestion of Drosophila DNA provided by the BDGP from the			
	isogenic strain Y2; cn bw sp, the same strain used for the BDGP's			
	pl and EST libraries. A more detailed description of the library			
	and how to order individual BAC clones, the entire library, or			
	filters for hybridization from the BACPAC Resource Center can be			
	found at http://bacpac.med.buffalo.edu/drosophila_bac.htm .			

ORIGIN

Query Match	3.1%;	Score 39.4;	DB 29;	Length 995;
Similarity	33.7%;	Pred. No. 12;		
Best Local				
Matches	55;	Conservative 43;	Mismatches 65;	Indels 0;
				Gaps 0.

[illegible]

RESULT 12					
BF341287/c					
LOCUS	BF341287	859 bp	mRNA	linear	EST 22-NOV-2000
DEFINITION	6020133222P1 NCI CGAP_Brn64 Homo sapiens cDNA clone IMAGE:4149244				
	5', mRNA sequence.				
ACCESSION	BF341287				
VERSION	BF341287.1	GI:11287778			
KEYWORDS	EST.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				

REFERENCE Mammalia; Euteria; Primates; Carnivora; Rodentia; Insectivora; Homo.
AUTHORS 1 (bases 1 to 859)
TITLE NIH-MGC <http://mgc.ncl.nih.gov/>.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgaabs-r@mail.nih.gov
Tissue Procurement: David N. Louis, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:

ORIGIN

QY 805 CTTGGGTCGTGACCATCCGGACTCTCTCTGGAAAAGGCGCTAAATGACAACTGAT 864
 Db 270 CTTGGGGGCGATGGCCATCTGCTCCCTGCACACTTTGACAAAGACAGACGTGAGACCGCAT 211
 QY 865 TTCTCTGTAATTCCTTGTGTGCGAGATGTGCGAG 901
 Db 210 TTTTGCTAAGGTACCTCTCTGTGAGGCTCTGCGAG 174

RESULT 9					
CNS009WA/C					
LOCUS	CNS009WA	787 bp	DNA	linear	GSS 03-JUN-1999
DEFINITION	Drosophila melanogaster genome survey sequence 17 end of Bac # BAC2004 of RPl-98 library from Drosophila melanogaster (fruit				

COMMENT

FEATURES

```
source
1. 787
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACR20K04"
/clone_1ib="RPCT-98"
/note="end : 77"
```

ORIGIN

Query Match	3.1%	Score 39.4	DB 29	length 787
Best Local Similarity	45.5%	Pred. No. 11		
Matches	76	Conservative	15	Mismatches 76; Indels 0; Gaps 0
QY	652	CGAGTTGCGCCCTTGTGCTCTCTTTGTGCAGCGTGA	CTTGTGATTCCTCGCGGACGAAAA	711
DB	334	CTACTGCTCCTTGTCTGCTTTTGTGTTAA	MMKMKMKMCTTGGATATCCKTKVAGABABAC	275
QY	712	GCAGATCCCTTCTGCTGCTGCTGCGACGGTTGAGCTT	KACCAATGCTCTTTGSGTAAATCC	771
DB	274	AATTGTTKTTTTCAWCGRCCTAAGGATAATTCT	TAATTTTMTGTTTAATTT	215
QY	772	AGGTAGGCCCTAATTGGCGGCTGCTAATCTCTT	GGSTGCTGTA	818
DB	214	TAATTTCGCACATGTTAATCGAATTTTTTTA	TGTCGCTGTA	168

RESULT 10
DVA0000000

LOCUS	885 bp	RNA	linear	EST 15-MAY-2003
DEFINITION	EX425603	Homo sapiens	NEUROBLASTOMA	Homo sapiens cDNA clone
	EX425603	U08B022A07	3-PRIME,	mRNA sequence.

ACCESSION	BX425603
VERSION	BX425603.1
	GI:30770486

SOURCE	Homo sapiens	(human
ORGANISM	Homo sapiens	

REFERENCE
AUTHORS
TITLE
JOURNAL

1 (bases 1 to 885)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
Li, W.B., Gruber, C., Jeejee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)

ORIGIN

QY 805 CTGAGGCTGTATACCAATCCGGACTCTTCTTGGGAAGCGTGTAAATACAACTGAT 864

Db 321 CTGGGGCAATGGCCATCTGCTCTGCAACCTTTGCAAAAGACAGACGTATAGCCGAT 264

QY 865 TTCTCCTGTATCTCCTGTATGTCGAGATATGAG 901

Db 261 TTGTGCTAAGGTCACCTTCTGTGCGAGCCTCTGAG 225

REFERENCE
AUTHORS
TITLE
JOURNAL

1 (pages 1 to 718)
Abdrakhmanov, I., Lodgjin, D., Gerolt, P., Arakawa, H., Law, A.,
Plachy, V., Korn, B. and Buerstedde, J. M.
A large database of chicken bursal ESTs as a resource for the
analysis of vertebrate gene function
Genome Res. 10 (12), 2062-2069 (2000)

COMMENT

FEATURES

```

/organism="Gallus gallus"
/mol_type="rRNA"
/strain="CB"
/dn_xref="Ectaxon.9031"
/clone="217r1"
/tissue_type="Bursa of Fabricius"
/cell_type="Bursal lymphocyte"
/dev_stage="2-3 weeks old"

```

Qy

Qy	Db	Qy	Db
805	117	865	57
CTTGGGCTCTGTTGAACATCCGGTACTCTTTTGGGAAGGCGCAAAAGCAACTGAT	CTGGGGGCGAGTTGGCCATCTGCTCCTGACCTTTTGGCAACAAGCAGACGTGAGACCCGAT	TTCCTCTGATATCTCCTTGTGTGTGGCCAGTATGTCAG	TTTTGCTAAGGTCAACCTTCTTGTCTGTGCAACCTCTGTGAG
864	58	901	21

RESULT 7	CC684922	743 bp	DNA	linear	GSS 19-JUN-2001
LOCUS	CC684922				
DEFINITION	CC684922				
ACCESSION	CC684922				
VERSION	CC684922.1				
KEYWORDS	GSS.				
SOURCE	zea mays				
ORGANISM	zea mays				

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
1 (Pass 1 to 743)	Whitlaw,C.A., Quackenbush,J., Van Aken,S., Uterback,T., Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T., Clerk,R.W., Nurnberg,A., Robbins,D. and Lakey,N.	Consortium for Maize Genomics	Unpublished (2002)	
Other_GSSs: CGUX55TV				

FEATURES
SOURCE

```

FEATURES
source
Location/Qualifiers
1..743
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMB0397J14"
/size_1kb="2M 0.7 1.5 KB"
/notes="Vector: pBCKS-1; Site 1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"

```

ORIGIN

Query Match	3.1%	Score 39.4	DB 29	Length 743
Best Local Similarity	53.6%	Pred. No. 11		
Matches 82	Conservative		Mismatches 71	Indels 0
				Gaps 0

QY

QY	591	TATTCGGCGGCTCTCACCGGAGTGGCGATCGCAGAGTTATTCCTTTGAATTAAAGGCC	65
Db	535	GTTCGCGGGCGGCACCCGGATCCAGACACGCTTCGCGAGCGCCTCTAAAGCTTCAGGTC	594
QY	651	TCGAGTTGCGCCCTTGCTCTCTCTTATGACAGC	683
Db	595	ACGAGAGGCGGCTTTGCTCTCTCCTGTCGCGC	627

[illegible]

```
1..1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="C83D1072YF05"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_1ib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo (dT)
```

ORIGIN

primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pGMSPORT 6 vector. Library was normalized.

	Query March	3.28; Score 40.8; DB 13; Length 1201;	
	Best Local Similarity	3.0%; Pred. No. 5.2;	
	Matches 17; Conservative 191; Mismatches 362; Indels 0; Gaps 0;		
Oy	397 CACATTGCTGGTGAACATCCGCCAGCATTCGTCTATGGCTTTCAATCCCCGCGCATGTGCT	456	
Dd	406 CKMMNNNNKKKKCKMKNKMKNN	465	
Oy	457 CAATAAACCATTGCCCGCTTCTATTCGGTTTTTCGGCCTTATGACGAAAGCCTACGCAGT	516	
Dd	466 MNNTNNNNNNNCHKANNNKNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNK	525	
Oy	517 CACTGCCGCCAGGCCGCGAGCTGTGGGGTGCGCATTAATCATGCAAATPAGGTT	576	
Dd	526 MMNNNCNMAMANNKNNKNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNKKKKK	585	
Oy	577 TCACTCTACTGGGTATTCGGCGCTCTACCGAGTGCGAGATCGCAGAATTGATTCCTT	636	
Dd	586 NNN	645	
Oy	637 TGAATTPAAGGCTCGAGTCCGCCCTTGGCTCTCTTTGTACGGTGACTTGATTC	696	
Dd	646 KNN	705	
Oy	697 CTGCCGAACGAAAAAGCAGATCCCTCTCTGCTGCTCGCAGGTTGAGCTTCAACATTCC	756	
Dd	706 NKNNKTTHKKKKNAKKNNNTNN	765	
Oy	757 TCTTGATGTAATTCAGAGTCAGGCGCTATTGCGAGCGCTGTGATCTTCTTGAGGCTGTT	816	
Dd	766 KNN	825	
Oy	817 GAACCATCCGGTACTTCTTTTGGGAAAAGGCTCTAAATGACAATGATTTTCTCTGATTT	876	
Dd	826 KKCMCKMMNMCKCMCKMKNN	885	
Oy	877 CTCCTGTGTTGCGAGTATGCGAGTCAATTAATTTGGCCTCCGGGCGGTTCCGTTCTTA	936	
Dd	886 MMKNN	945	
Oy	937 ATCCTTAAGCCCCTPAACGTAATCAATT 966		
Dd	946 MNNECVNMKKNNKCNCKKVNMKKCARKKK 975		

RESULT 5
 BX540413/c
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT
 692 bp mRNA linear EST 28-JUL-2001
 BX5500413 AGENAE Gallus gallus multi-tissues library (ycab) Gallus
 gallus cDNA clone gcab0001b.d.10 5prim, mRNA sequence.
 BX540413
 BX540413
 BX540413.1 GI:33294986
 EST.
 Gallus gallus (chicken)
 Gallus gallus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
 Phasianinae; Gallus.
 1 (bases 1 to 692).
 Herault,F., Le Meuth-Metzinger,V., Desert,C., Retout,F., Piumi,F.,
 Klopp,C. and Donati,M.
 Construction and primary characterization of chicken normalized
 multi-tissue cDNA libraries
 Unpublished (2003)
 Contact: Donati M
 INRA, UMR INRA-ENSAR Genetique Animale
 65, rue de Saint-Brieuc, RENNES cedex, 35042, FRANCE
 Tel: +33 (0) 2.23.48.54.63

Email: Raf.Podowski@agr.ki.se
Class: NotI site
Location/Qualifiers

FEATURES

Source

1. 473
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone_lib="Human NotI clones"

ORIGIN

Query Match 4.3%; Score 54.6; DB 28; Length 473;
Best Local Similarity 56.2%; Pred. No. 0.0036;
Matches 99; Conservative 0; Mismatches 77; Indels 0; Gaps 0;

QY 293 TGGGACGCCCCACTGTTTCCGGCCGATTTTCGGGGGCTCCACGAATGCTGTCATC 352
DB 283 TGGGAGGGCCAGGGGCTCTGGCCATGCTTCGCGGGGCGCCGACATTTGTGGCCATC 342
QY 353 GCCCTCGTTGTGGCGCAGCGCCCTGGCGCCATCGGCTCACCACTTGTGTGAAC 412
DB 343 GGCATGCTTAAAGGCGGACCAACCTGCTGATCTGTCATCCACACACTGCTGTACT 402
QY 413 TTCCGCCACGATTCATTCGCTTTTCATCCCGCTGACATGTCGAAAACCCCAT 468
DB 403 TCGCAGACCTGCTTACGCGCTTGTCATGCGCCCGCTGCTTACAGCANCCTT 458

RESULT 2
BX356664 1201 bp mRNA linear EST 05-MAY-2003
LOCUS BX356664 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
DEFINITION clone CS0D1015YB03 3-PRIME, mRNA sequence.
ACCESSION BX356664
VERSION BX356664
KEYWORDS BX356664.1 GI.30378083
SOURCE EST.
ORGANISM Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 1201)
AUTHORS Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr

Library was constructed by Life Technologies, a division of
Invitrogen. Contact: Feng Liang Email: fliang@lifetech.com URL :
http://fulllength.invitrogen.com/Invitrogen Corporation 1600
Paradey Avenue Genoscope sequence ID : CS0D1015CA02NP1.

FEATURES

Source

1. 1201
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1015YB03"
/issue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 3.5%; Score 45; DB 13; Length 1201;
Best Local Similarity 12.9%; Pred. No. 0.32; Indels 0; Gaps 0;
Matches 47; Conservative 168; Mismatches 148; Indels 0; Gaps 0;
QY 307 GTTTCGCGCCTGATTTTCGGGGCTCCACGAATGCTGTCATCGCCCTGTTGGG 366
DB 745 STBST 804

QY 367 CGACGCCCCCTGGGGCCATCGCCCTACACCACTTGTGTAACCTCCGCCATAT 426
DB 805 SBTCTSSSSSSSSSBTTTSTSTKSBTTTBSBTTSSSTSTSTSTSTSTSTST 864
QY 427 CTATGCTTTTATCCCGCTGCATGTGTCAAAAACCCATTCGCCCTTTCTATCGT 486
DB 865 YSBSST 924
QY 487 TTTCGCGCTTATCGAAGACCTACGACTCATCGCGCCAGCCCGCAGCTGTGCGC 546
DB 925 BTSSSSSBTTTSSSSSSSBTTTSSSSSBTTTSSSSSBTTTSSSSSBTTTSS 984
QY 547 GTGGCAGCTATCTCAATGCAAAATAGCGTTCACTCTGATTCGCGGCTCTCAC 606
DB 985 VTSSSBTTST 1044
QY 607 CGAGTGCAGATCGACAGATGATTCCTTTGAAATTAAGGCGCTGACTGCCCTTTG 666
DB 1045 SSSSSSSSSSSSSSSSBTTTBTSTSTSTSTSTSTSTSTSTSTSTSTSTST 1104
QY 667 CTC 669
DB 1105 TTS 1107

RESULT 3
BX376097 1201 bp mRNA linear EST 08-MAY-2003
LOCUS BX376097 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens
DEFINITION cDNA clone CS0DC022YM12 5-PRIME, mRNA sequence.
ACCESSION BX376097
VERSION BX376097.1 GI.30434756
KEYWORDS EST.
SOURCE EST.
ORGANISM Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 1201)
AUTHORS Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr

Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 2866.f
Contact: Feng Liang Email: fliang@lifetech.com URL :
http://fulllength.invitrogen.com/Invitrogen Corporation 1600
Paradey Avenue Genoscope sequence ID : CS0DC022B06P1.

FEATURES

Source

1. 1201
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DC022YM12"
/issue_type="NEUROBLASTOMA COT 25-NORMALIZED"
/clone_lib="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 3.3%; Score 41.8; DB 13; Length 1201;
Best Local Similarity 9.1%; Pred. No. 2.7;
Matches 55; Conservative 238; Mismatches 310; Indels 2; Gaps 1;
QY 593 TTTCGCGCTTCACCGAGTGGCGATCGCAAGTGTATCTCTTTGAATTAAGGCTC 652
DB 441 TTTTNTTNTTNTTMMTKKNTTKKNCCKKKKNNMMCCNNCKKKTKTKKKKKNNKKNN 500
QY 653 GAGTTGCGCCTTGTCTCTCTTGTCAAGCTGACTTGTGATTCCTGCCGAAGAAAAG 712

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 17, 2004, 06:48:18 ; Search time 415.78 Seconds
(without alignments)
9221.792 Million cell updates/sec

Title: US-10-608-504-6

Perfect score: 1271

Sequence: 1 ggcggatcaatggaatctag.....agccttaacgacagcgc 1271

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
EST:
1: em_estba:*
2: em_estum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gse_hum:*
18: em_gse_huv:*
19: em_gse_pin:*
20: em_gse_vit:*
21: em_gse_fun:*
22: em_gse_mam:*
23: em_gse_mus:*
24: em_gse_pro:*
25: em_gse_rtd:*
26: em_gse_phg:*
27: em_gse_vrl:*
28: gb_gse1:*
29: gb_gse2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	54.6	4.3	473	28	AQ936633 HSJ41-905
2	45	3.5	1201	13	BX356664 BX356664
3	41.8	3.3	1201	13	BX376097 BX376097
4	40.8	3.2	1201	13	BX381961 BX381961

RESULT 1	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	MEDLINE	PIRME	COMMENT
AQ936633	HSJ41-905	Human NotI clones Homo sapiens genomic, genomic survey	AQ936633	GI:7213011			Homo sapiens (human)	1 (bases 1 to 473)	Zabarovskiy, E.R., Gizatullin, R., Podowski, R.M., Zabarovskaya, V.V., Xie, L., Muravenko, O.V., Kozlyev, S., Petrenko, L., Skobeleva, N., Li, J., Protodopov, A., Kashuba, V., Erber, I., Winberg, G. and Wahlstedt, C.	NotI clones in the analysis of the human genome	Nucleic Acids Res. 28 (7), 1635-1639 (2000)	10710430	Contact: Podowski RM Center for Genomics Research Karolinska Institute 17177 Stockholm, Sweden Tel: +46-8-728-6372 Fax: +46-8-337983	
5	39.4	3.1	692	13	BX540413									BX540413 BX540413
6	39.4	3.1	718	9	AJ397466									AJ397466 AJ397466
7	39.4	3.1	743	29	CC684922									CC684922 OGUXX55TH
8	39.4	3.1	756	13	BU448643									BU448643 603211326
9	39.4	3.1	787	29	CNS009WA									ALU54013 Drosophila
10	39.4	3.1	885	13	BX425603									ALU54013 Drosophila
11	39.4	3.1	885	13	BX425603									ALU54013 Drosophila
12	39.4	3.1	885	13	BX425603									ALU54013 Drosophila
13	39.4	3.1	885	13	BX425603									ALU54013 Drosophila
14	39.4	3.1	885	13	BX425603									ALU54013 Drosophila
15	39.4	3.1	885	13	BX425603									ALU54013 Drosophila
16	39.4	3.1	885	13	BX425603									ALU54013 Drosophila
17	39.4	3.1	885	13	BX425603									ALU54013 Drosophila
18	39.4	3.1	885	13	BX425603									ALU54013 Drosophila
19	39.4	3.1	885	13	BX425603									ALU54013 Drosophila
20	39.4	3.1	885	13	BX425603									ALU54013 Drosophila
21	39.4	3.1	885	13	BX425603									ALU54013 Drosophila
22	39.4	3.1	885	13	BX425603									ALU54013 Drosophila
23	39.4	3.1	885	13	BX425603									ALU54013 Drosophila
24	39.4	3.1	885	13	BX425603									ALU54013 Drosophila
25	39.4	3.1	885	13	BX425603									ALU54013 Drosophila
26	39.4	3.1	885	13	BX425603									ALU54013 Drosophila
27	39.4	3.1	885	13	BX425603									ALU54013 Drosophila
28	39.4	3.1	885	13	BX425603									ALU54013 Drosophila
29	39.4	3.1	885	13	BX425603									ALU54013 Drosophila
30	39.4	3.1	885	13	BX425603									ALU54013 Drosophila
31	39.4	3.1	885	13	BX425603									ALU54013 Drosophila
32	39.4	3.1	885	13	BX425603									ALU54013 Drosophila
33	39.4	3.1	885	13	BX425603									ALU54013 Drosophila
34	39.4	3.1	885	13	BX425603									ALU54013 Drosophila
35	39.4	3.1	885	13	BX425603									ALU54013 Drosophila
36	39.4	3.1	885	13	BX425603									ALU54013 Drosophila
37	39.4	3.1	885	13	BX425603									ALU54013 Drosophila
38	39.4	3.1	885	13	BX425603									ALU54013 Drosophila
39	39.4	3.1	885	13	BX425603									ALU54013 Drosophila
40	39.4	3.1	885	13	BX425603									ALU54013 Drosophila
41	39.4	3.1	885	13	BX425603									ALU54013 Drosophila
42	39.4	3.1	885	13	BX425603									ALU54013 Drosophila
43	39.4	3.1	885	13	BX425603									ALU54013 Drosophila
44	39.4	3.1	885	13	BX425603									ALU54013 Drosophila
45	39.4	3.1	885	13	BX425603									ALU54013 Drosophila

ALIGNMENTS

Tue Apr 20 06:47:23 2004

us-10-608-504-6.rnpn

Page 6

[illegible]

Search completed: April 17, 2004, 20:13:02
Job time : 148.874 secs

```
RESULT 12
US-10-021-698A-716
; Sequence 716, Application US/10021698A
; GENERAL INFORMATION:
; APPLICANT: KEITH, TIM
; APPLICANT: LITTLE, RANDALL
; APPLICANT: VAN EERDENEGH, PAUL
; APPLICANT: DUPUIS, JOSEF
; APPLICANT: DEL MASTRO, RICHARD
; APPLICANT: SIMON, JASON
; APPLICANT: ALLEN, KRISTINA
; APPLICANT: PANDIT, SUNIL
; TITLE OF INVENTION: NUCLEOTIDE AND AMINO ACID SEQUENCES RELATING TO
; TITLE OF INVENTION: RESPIRATORY DISEASES AND OBESITY
; FILE REFERENCE: 29/6-404US1
; CURRENT APPLICATION NUMBER: US/10/021,698A
; CURRENT FILING DATE: 2001-10-22
; PRIOR APPLICATION NUMBER: 60/211,749
; PRIOR FILING DATE: 2000-06-14
; NUMBER OF SEQ ID NOS: 6160
; SOFTWARE: PasteSeq for Windows Version 4.0
; SEQ ID NO 716
; LENGTH: 173233
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: modified_base
; LOCATION: (88987)..(89086)
; OTHER INFORMATION: a, t, c or g
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (167286)..(167385)
; OTHER INFORMATION: a, t, c or g
US-10-021-698A-716

Query Match
Best Local Similarity 55.1%; Pred. No. 39;
Matches 65; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

QY 1052 AGACCTAAGCTTGTGCTCATTTGCGCTGACGATTAAGTGGGCGCATCTTCTGCG 1111
DB 84488 AGAATTAAGCTTGTGCTCATTTGCGCTGACGATTAAGTGGGCGCATCTTCTGCG 84547
QY 1112 GTGAGCAGCACTGTTGTGAGCGTTGGCGCTGGACCATCGTTTGTGAGCTGTGAA 1169
DB 84548 GTAACTCAACACCTGTGCTGTTGGCGCTGTACCATCGTGTGATTTGGAGATGAA 84605

RESULT 13
US-60-548-091-19934
; Sequence 19934, Application US/60548091
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: STROKE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C1001506
; CURRENT APPLICATION NUMBER: US/60/548,091
; CURRENT FILING DATE: 2004-02-27
; NUMBER OF SEQ ID NOS: 24433
; SOFTWARE: PasteSeq for Windows Version 4.0
; SEQ ID NO 19934
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-60-548-091-19934

Query Match
Best Local Similarity 52.7%; Pred. No. 0.63;
Matches 69; Conservative 1; Mismatches 61; Indels 0; Gaps 0;

QY 342 TGCTGTATCGCCCTGCTGTGTGTGGGCGACGCGCCCTGTGGCGCATCGGCTTACACAT 401
```

```
DB 46 TGCTGTATCGGCGCTCCCTGATGGCTCGCTGGCTCCGAGATGACAGGCGGCTTCT 105
QY 402 TGCTGTATCGACTTCCCGCACGATTTATGCGTTTATCCCGCTGCATGTGTCAAAA 461
DB 106 GCTTCAAGAGACCCCGCACATTTCACTTCACTGAGCGTTCCGCTGTGGATTTAC 165
QY 462 ACCCATTTGCC 472
DB 166 TGCCCTTGCC 176

RESULT 14
US-60-548-091-19938
; Sequence 19938, Application US/60548091
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: STROKE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C1001506
; CURRENT APPLICATION NUMBER: US/60/548,091
; CURRENT FILING DATE: 2004-02-27
; NUMBER OF SEQ ID NOS: 24433
; SOFTWARE: PasteSeq for Windows Version 4.0
; SEQ ID NO 19938
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-60-548-091-19938

Query Match
Best Local Similarity 52.6%; Score 33; DB 7; Length 201;
Matches 69; Conservative 1; Mismatches 61; Indels 0; Gaps 0;

QY 342 TGCTGTATCGCCCTGCTGTGTGTGGGCGACGCGCCCTGTGGCGCATCGCTTACACAT 401
DB 49 TGCTGTATCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 108
QY 402 TGCTGTATCGACTTCCCGCACGATTTATGCGTTTATCCCGCTGCATGTGTCAAAA 461
DB 109 GCTTCAAGAGACCCCGCACATTTCACTTCACTGAGCGTTCCGCTGTGGATTTAC 168
QY 462 ACCCATTTGCC 472
DB 169 TGCCCTTGCC 179

RESULT 15
US-10-767-701-9054
; Sequence 9054, Application US/10767701
; GENERAL INFORMATION:
; APPLICANT: KOVALIC, David K.
; APPLICANT: Zhou, Yihua
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(5335)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 9054
; LENGTH: 1028
; TYPE: DNA
; ORGANISM: Sorghum bicolor
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-CLUS73927_1
US-10-767-701-9054

Query Match
Best Local Similarity 54.0%; Pred. No. 2.1;
Matches 67; Conservative 0; Mismatches 57; Indels 0; Gaps 0;

QY 709 AAAGCATGCTCTGCTGTGTGTGGGCGACGCTTGAAGTTGACATGCTTGTGTAT 768
```



```
Query Match          2.7%; Score 34; DB 6; Length 561;
Best Local Similarity 54.9%; Pred. No. 0.56;
Matches 67; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

QY 302 CCACGTGTTTCCGGCCTGATTTTCGGGGCTCCACGAAATGCTGATCGCCCTGTT 361
DB 333 CCACGAGAGCGGACCGCTGAGCTTCTGAGCCGAGAGCCGATCTGAGGCTGCTGTC 274
QY 362 GTGGGGGAGCGCCCTGGGGCCGATCGGCTCAGCAATTCGTGTAACTTCGGCCAC 421
DB 273 ACCAGGCGGACGCGCTGCGGCTTCTCTCCAGCAGCCTGTGTAGTAGGTAAGCGCC 214
QY 422 GT 423
DB 213 CT 212

RESULT 9
US-10-765-790-62
; Sequence 62; Application US/10765790
; GENERAL INFORMATION:
; APPLICANT: Bayer Healthcare LLC
; APPLICANT: Beard, Chris
; APPLICANT: Burgess, Chris
; APPLICANT: Gannon, Allison
; APPLICANT: Harvey, Jeanne
; APPLICANT: Lechner, John F.
; APPLICANT: Li, Zheng
; TITLE OF INVENTION: Identification and Verification of Methylation Marker Sequences
; FILE REFERENCE: 1657/2035
; CURRENT FILING DATE: 2004-01-27
; PRIOR FILING DATE: 2003-12-16
; PRIOR APPLICATION NUMBER: US 10/737,082
; NUMBER OF SEQ ID NOS: 300
; SOFTWARE: Patent version 3.2
; SEQ ID NO 62
; LENGTH: 23490
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-765-790-62

Query Match          2.7%; Score 33.8; DB 6; Length 23490;
Best Local Similarity 49.7%; Pred. No. 7;
Matches 86; Conservative 0; Mismatches 87; Indels 0; Gaps 0;

QY 301 CCCACTGTTTCCGGCCTGATTTTCGGGGCTCCACGAAATGCTGATCGCCCTGTT 360
DB 21210 CTGGGTCCTCTCCCTCCCTGATCTGTGTTCCACCTTCTGCTCTCAGGCTTGGA 21269
QY 361 TGTGGGGGAGCGCCCTGGGGCCGATCGGCTCAGCAATTCGTGTAACTTCGGCA 420
DB 21270 GAGAGTGGGGGCTGGGCTGAGCTGAGCGGCTATGCGCTCTGAGGCGCTGACT 21329
QY 421 CGATTCTATGCGTTTTCATTCCTCGCTGATGTGTGTAACCCCATTTGCC 473
DB 21330 CCTCTCATGCGCTTCTCTTGTGCTGAGCGGGCGGTATACCACTTGCCC 21382

RESULT 10
US-60-548-091-5686
; Sequence 5686; Application US/60548091
; GENERAL INFORMATION:
; APPLICANT: CARBELL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: STROKE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C1001506
; CURRENT FILING DATE: US/60/548,091
; NUMBER OF SEQ ID NOS: 24433
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5686
```

```
; LENGTH: 24477
; TYPE: DNA
; ORGANISM: Homo sapiens
US-60-548-091-5686

Query Match          2.7%; Score 33.8; DB 7; Length 24477;
Best Local Similarity 51.9%; Pred. No. 7.1;
Matches 68; Conservative 3; Mismatches 60; Indels 0; Gaps 0;

QY 342 TGCTGTGATCGCGCTGTTGTGGGGCAGCGCCCTGGGGCCGATCGGCGCTCAGCAAT 401
DB 17603 TGCTGTGATCGCGCTGTTGTGGGGCAGCGCCCTGGGGCCGATCGGCGCTCAGCAAT 17662
QY 402 TGCTGTGATCGCGCTGTTGTGGGGCAGCGCCCTGGGGCCGATCGGCGCTCAGCAAT 461
DB 17663 GCTTCAGAGCAGCGCGGACATTTCAATTCAGTGGCGTTCGCGTGGGATTTTACC 17722
QY 462 ACCCGATTGCC 472
DB 17723 TGCCCTGGCC 17733

RESULT 11
US-09-804-291A-500
; Sequence 500; Application US/09804291A
; GENERAL INFORMATION:
; APPLICANT: ZOZULA, SERGEY
; TITLE OF INVENTION: HUMAN OLFACTORY RECEPTORS AND GENES ENCODING SAME
; FILE REFERENCE: 100317.54287US
; CURRENT FILING DATE: US/09/804,291A
; CURRENT FILING DATE: 2001-03-13
; PRIOR FILING DATE: 2000-03-13
; PRIOR FILING DATE: 2000-03-13
; PRIOR FILING DATE: 2000-03-24
; PRIOR FILING DATE: 2000-03-24
; PRIOR FILING DATE: 2000-04-14
; PRIOR FILING DATE: 2000-04-14
; PRIOR FILING DATE: 2000-04-24
; PRIOR FILING DATE: 2000-04-24
; PRIOR FILING DATE: 2000-05-26
; PRIOR FILING DATE: 2000-05-26
; PRIOR FILING DATE: 2000-06-23
; PRIOR FILING DATE: 2000-06-23
; PRIOR FILING DATE: 2000-08-16
; PRIOR FILING DATE: 2000-08-16
; PRIOR FILING DATE: 2000-09-07
; PRIOR FILING DATE: 2000-09-07
; PRIOR APPLICATION NUMBER: 60/207,702
; PRIOR APPLICATION NUMBER: 60/207,702
; PRIOR APPLICATION NUMBER: 60/213,849
; PRIOR APPLICATION NUMBER: 60/213,849
; PRIOR APPLICATION NUMBER: 60/226,534
; PRIOR APPLICATION NUMBER: 60/226,534
; PRIOR APPLICATION NUMBER: 60/230,732
; PRIOR APPLICATION NUMBER: 60/230,732
; PRIOR FILING DATE: 2001-02-07
; PRIOR FILING DATE: 2001-02-07
; NUMBER OF SEQ ID NOS: 531
; SOFTWARE: Patent Ver. 3.2
; SEQ ID NO 500
; LENGTH: 930
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-804-291A-500

Query Match          2.6%; Score 33.2; DB 5; Length 930;
Best Local Similarity 54.0%; Pred. No. 1.4;
Matches 68; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

QY 567 AATAGCGTTTCACTCCCTACCTGCTATTCGGCGGCTCCACCGAGTGGCGATGGCAGAGT 626
DB 509 ACATTTCTAATCATTATCTGTACTATGAGGCTCTGAGAGCTTGGCTCTCAGACA 568
QY 627 TGATTCCTTTGAATTAAGGCGCTCGAGTTGGCGCTTGTCTCTCTTTGACGCTGA 686
DB 569 CAGGCTCTTGAAGTATGATCTCTTGGCGGCTGTGACTCTATGTTACTCTGG 628
QY 687 CTTTGG 692
DB 629 TGCTGG 634
```

```
; Sequence 983, Application US/10100683
; GENERAL INFORMATION:
; APPLICANT: Rosen, et al.
; TITLE OF INVENTION: Human Secreted Proteins
; FILE REFERENCE: PS900
; CURRENT APPLICATION NUMBER: US/10/100,683
; CURRENT FILING DATE: 2002-03-19
; PRIOR APPLICATION NUMBER: US 60/040,162
; PRIOR FILING DATE: 1997-03-07
; PRIOR APPLICATION NUMBER: US 60/043,576
; PRIOR FILING DATE: 1997-04-11
; PRIOR APPLICATION NUMBER: US 60/047,601
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: US 60/056,845
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: US 60/043,314
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: US 60/056,664
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: US 60/047,632
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: US 60/056,892
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 13468
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 983
; LENGTH: 3173
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-100-683-983

Query Match      2.7%; Score 34.6; DB 6; Length 3173;
Best Local Similarity 75.4%; Pred. No. 1;
Matches 43; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

Cy 344 CTGGTCATGCGCCCTGTGTGGGCGGAGCGCCCTGGGCGGCATGCGCTCACACA 400
Db 2360 CAGCTCCCGCCGACGTGTGGGCTCAGGGCTCTCGGCGCCATGCGTCTCACACA 2304

RESULT 6
US-10-100-683-11511
; Sequence 11511, Application US/10100683
; GENERAL INFORMATION:
; APPLICANT: Rosen, et al.
; TITLE OF INVENTION: Human Secreted Proteins
; FILE REFERENCE: PS900
; CURRENT APPLICATION NUMBER: US/10/100,683
; CURRENT FILING DATE: 2002-03-19
; PRIOR APPLICATION NUMBER: US 60/040,162
; PRIOR FILING DATE: 1997-03-07
; PRIOR APPLICATION NUMBER: US 60/043,576
; PRIOR FILING DATE: 1997-04-11
; PRIOR APPLICATION NUMBER: US 60/047,601
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: US 60/056,845
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: US 60/043,580
; PRIOR FILING DATE: 1997-04-11
; PRIOR APPLICATION NUMBER: US 60/047,599
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: US 60/056,664
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: US 60/043,314
; PRIOR FILING DATE: 1997-04-11
; PRIOR APPLICATION NUMBER: US 60/047,632
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: US 60/047,632
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: US 60/056,892
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 13468
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 983
; LENGTH: 3173
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-100-683-983
```

```
; PRIOR FILING DATE: 1997-08-22
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 13468
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11511
; LENGTH: 7071
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-100-683-11511

Query Match      2.7%; Score 34.6; DB 6; Length 7071;
Best Local Similarity 75.4%; Pred. No. 1.7;
Matches 43; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

Cy 344 CTGGTCATGCGCCCTGTGTGGGCGGAGCGCCCTGGGCGGCATGCGCTCACACA 400
Db 785 CAGCTCCCGCCGACGTGTGGGCTCAGGGCTCTCGGCGCCATGCGTCTCACACA 841

RESULT 7
US-10-767-701-7545
; Sequence 7545, Application US/10767701
; GENERAL INFORMATION:
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(5353)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 7545
; LENGTH: 666
; TYPE: DNA
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-CLUS92927_1
US-10-767-701-7545

Query Match      2.7%; Score 34.4; DB 6; Length 666;
Best Local Similarity 57.4%; Pred. No. 0.46;
Matches 62; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

Cy 315 GCCTGATTTTGGGGGCTCCACCGAATGCTGCTATCGCCCTGTTGGGGCGGCGC 374
Db 31 GCCCGACACACAGGGGGAGAGACCCGCTCATCAACCGCTTGGGAGCGGCGCGC 90

Cy 375 CCTGGGCGGCATGCGGCTCACACCAATGCTGTGAATTCCGCGACG 422
Db 91 CGCTGGCGTCTCGGCTGTGACGCGCGTGTCTGTGTGCGGTGCGCGAG 138

RESULT 8
US-10-767-701-744/c
; Sequence 744, Application US/10767701
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(5353)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 744
; LENGTH: 561
; TYPE: DNA
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-CLUS124346_1
US-10-767-701-744
```

NUMBER OF SEQUENCES: 7310
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: GENOME THERAPEUTICS CORPORATION
 STREET: 100 Beaver Street
 CITY: Waltham
 STATE: Massachusetts
 COUNTRY: USA
 ZIP: 02354
 COMPUTER READABLE FORM:
 MEDIUM TYPE: CD-ROM ISO9660
 COMPUTER: PC
 OPERATING SYSTEM: <Unknown>
 SOFTWARE: ASCII
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/10/417,884A
 FILING DATE: 17-Apr-2003
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/09/107,532
 FILING DATE: 30-Jun-1998
 APPLICATION NUMBER: 60/085,598
 FILING DATE: 14 May 1998
 APPLICATION NUMBER: 60/051,571
 FILING DATE: July 2, 1997
 ATTORNEY/AGENT INFORMATION:
 NAME: Arinello, Pamela Deneke
 REGISTRATION NUMBER: 40,489
 REFERENCE/DOCKET NUMBER: GTC-012
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (781)893-5007
 TELEFAX: (781)893-8277
 INFORMATION FOR SEQ ID NO: 1266:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 768 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: circular
 MOLECULE TYPE: DNA (genomic)
 HYPOTHEICAL: NO
 ANTI-SENSE: NO
 ORIGINAL SOURCE:
 ORGANISM: Enterococcus faecium
 FEATURE:
 NAME/KEY: misc feature
 LOCATION: (S) LOCATION 1..768
 SEQUENCE DESCRIPTION: SEQ ID NO: 1266:
 US-10-417-884A-1266
 Query Match 3.1%; Score 39.8; DB 6; Length 768;
 Best Local Similarity 49.3%; Pred. No. 0.0075;
 Matches 104; Conservative 0; Mismatches 107; Indels 0; Gaps 0;
 QY 207 TAAAACTCCCTTGGCTGAGGTTGGGCGATGACCCGATGGTATGGCTTTGGCTCT 266
 DB 92 TCAAGACAGTTACCTACCGTTTCGGTTATATCGGATTTGACTGATTTGGATCG 151
 QY 267 TGGTATTCAATACGCTACGAATGGGCGAGCCCACTGTTTCGGGCTGATTTTCG 326
 DB 152 TTGGGAAGCTGGCGGATTTCATCCACTAGTCGTACAGTTGATGTCCTGGCTGTATG 211
 QY 327 GGGGCTCACCGAAATCTGTATGCGCTTGTGTGGGCGAGGCGCCCTGGGCGCA 386
 DB 212 CTGGTTCTGCCAATTATACAGTCAGCATGCTTGTGTATACAGCCCATTTGCTTCA 271
 QY 387 TCGGCTCACCATGCTGTGTAATTCG 417
 DB 272 TCGTTTCTGACCTTTCTAGTCAATTCGG 302

RESULT 3
 US-10-767-701-5263
 ; Sequence 5263, Application US/10767701
 ; GENERAL INFORMATION:
 ; APPLICANT: Kovacic, David K.

APPLICANT: Zhou, Yihua
 TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
 TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
 FILE REFERENCE: 38-21(53535)B
 CURRENT APPLICATION NUMBER: US/10/767,701
 CURRENT FILING DATE: 2004-01-29
 NUMBER OF SEQ ID NOS: 63128
 SEQ ID NO 5263
 LENGTH: 619
 TYPE: DNA
 ORGANISM: Sorghum bicolor
 FEATURE:
 OTHER INFORMATION: Clone ID: SORBI-28MAY03-CLUS111974_1
 US-10-767-701-5263

Query Match 2.8%; Score 35.4; DB 6; Length 619;
 Best Local Similarity 54.1%; Pred. No. 0.2;
 Matches 72; Conservative 0; Mismatches 61; Indels 0; Gaps 0;
 QY 326 GCGGCTCCACCGAATGCTGTCATGCGCCCTGTTGGGCGCAGCGCCCTGGGCGCC 385
 DB 132 GCGGCTTCACTTGGCTGCTGCTGCGGCGGTGACGGGCAAGTGTGCGCTC 191
 QY 386 ATCGGCTCACCATTTGCTGGTGAATTCGCGCAGTATTCATGCTTTGATCCG 445
 DB 192 ATCTGCGGAGCGCGCTGCTGTGATCTTCAGCCCGTGTGTCGCGCGCATTCAG 251
 QY 446 CTGCATGGGTCA 458
 DB 252 GTGGCGCTGCTCA 264

RESULT 4
 US-10-767-701-10866
 ; Sequence 10866, Application US/10767701
 ; GENERAL INFORMATION:
 ; APPLICANT: Kovacic, David K.
 ; APPLICANT: Zhou, Yihua
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
 ; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
 ; FILE REFERENCE: 38-21(53535)B
 ; CURRENT APPLICATION NUMBER: US/10/767,701
 ; CURRENT FILING DATE: 2004-01-29
 ; NUMBER OF SEQ ID NOS: 63128
 ; SEQ ID NO 10866
 ; LENGTH: 1034
 ; TYPE: DNA
 ; ORGANISM: Sorghum bicolor
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: SORBI-28MAY03-CLUS4850_1
 US-10-767-701-10866

Query Match 2.7%; Score 34.6; DB 6; Length 1034;
 Best Local Similarity 49.7%; Pred. No. 0.52;
 Matches 88; Conservative 0; Mismatches 89; Indels 0; Gaps 0;

QY 341 ATGCGTGCATGCGCTGTTGTGGGCGAGCGCCCTGGGCGGCATGCGCTCACCA 400
 DB 151 ATGGCGGCGCTTACTAGGCGGAGGCTTGGCTGCGCACCGGCGGCGCCCT 210
 QY 401 TTGCTGTGAATCTCCGCGCATATTATGCGTTTATCTCCGCTGCATGTGTA 460
 DB 211 GTGGCGGTGTGTCGCGCATTTCTGCGCGCTTACGATGAGTCCCTGACGCTACCAAG 270
 QY 461 AACCCATGCGCGTTTCTATTGCTTGTGCGCTTATTCAGAGCAAGCTACGAGTC 517
 DB 271 AAGGCTACAGCTTCTCCAGCGGCACTTACCGTACCGAGCGCAAGGCGCCCTC 327

RESULT 5
 US-10-100-683-983/c

ENTREPRENEUR & ON D'INNOVATION ET D'INVESTISSEMENT

Tue Apr 20 06:47:23 2004

us-10-608-504-6.rmpb

Page 10

PRIOR FILING DATE: 2001-05-30
 PRIOR APPLICATION NUMBER: JP 2001-272697
 PRIOR FILING DATE: 2001-08-02
 NUMBER OF SEQ ID NOS: 15109
 SEQ ID NO 6067
 LENGTH: 594
 TYPE: DNA
 ORGANISM: Streptomyces avermitilis
 FEATURES:
 NAME/KEY: CDS
 LOCATION: (1)..(594)
 US-10-156-761-6067

Query Match 3.4%; Score 43; DB 14; Length 594;
 Best Local Similarity 50.7%; Pred. No. 0.0029;
 Matches 103; Conservative 0; Mismatches 100; Indels 0; Gaps 0;

QY 324 TCAGGAGCTCCACCAAAATGCTGTCATCGCCCTGTTGGGGGACGCGCCCTGGGGC 383
 DB 203 TCACGGCTGCGCCATCGCGAGGCTGCGCATGCTGCGACCGCGCTCGGGTGG 262
 QY 384 CCATCGCGCTCCACCAATGCTGTCATCGCCCTGTTGGGGGACGCGCCCTGGGGC 443
 DB 263 GCAAGTCCGACGACGATCGCGGATCGCTGCTTCTTCTGGGCTACTGCTCA 322
 QY 444 CGCTGATGTGTCAAAAACCCCATTTGCCCTTTCTATTGCTTTGCGGCTTATCGACG 503
 DB 323 CATTGCGGAGTGTGTGAAGCGCGGTGTCGACCGCTTGGGGTGGCGCTCG 382
 QY 504 AACCTAGCAGTCACTGGGCGC 526
 DB 383 CCGCGACACCTGTCTCATCGCC 405

RESULT 14

US-10-156-761-1/C
 Sequence 1, Application US/10156761
 Publication No. US20030119018A1
 GENERAL INFORMATION:
 APPLICANT: OKURA, SATOSHI
 APPLICANT: IKEDA, HARUO
 APPLICANT: ISHIKAWA, JUN
 APPLICANT: HORIKAWA, HIROSHI
 APPLICANT: SHIBA, TADAYOSHI
 APPLICANT: SAKAKI, YOSHIYUKI
 APPLICANT: HATTORI, MASAHIRA
 TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
 FILE REFERENCE: 249-262
 CURRENT APPLICATION NUMBER: US/10/156,761
 PRIOR FILING DATE: 2002-05-29
 PRIOR APPLICATION NUMBER: JP 2001-204089
 PRIOR FILING DATE: 2001-05-30
 PRIOR APPLICATION NUMBER: JP 2001-272697
 PRIOR FILING DATE: 2001-08-02
 NUMBER OF SEQ ID NOS: 15109
 SEQ ID NO 1
 LENGTH: 9025608
 TYPE: DNA
 ORGANISM: Streptomyces avermitilis
 FEATURES:
 NAME/KEY: misc_feature
 LOCATION: (4187715)
 OTHER INFORMATION: a, t, c, g, other or unknown
 US-10-156-761-1

Query Match 3.4%; Score 43; DB 14; Length 9025608;
 Best Local Similarity 50.7%; Pred. No. 1.4;
 Matches 103; Conservative 0; Mismatches 100; Indels 0; Gaps 0;
 QY 324 TCAGGAGCTCCACCAAAATGCTGTCATCGCCCTGTTGGGGGACGCGCCCTGGGGC 383
 DB 7327430 TCACGGCTGCGCCATCGCGAGGCTGCTGCGCATGCTGCGACCGCGCTCGGCTGG 7327371

QY 384 CCATCGCGCTCCACCAATGCTGTCATCGCCCTGTTGGGGGACGCGCCCTGGGGC 443
 DB 7327370 GCAAGTCCGACGACGATCGCGGATCGCTGCTTCTTCTGGGCTACTGCTCA 7327311
 QY 444 CGCTGATGTGTCAAAAACCCCATTTGCCCTTTCTATTGCTTTGCGGCTTATCGACG 503
 DB 7327310 CATTGCGGAGTGTGTGAAGCGCGGTGTCGACCGCTTGGGGTGGCGCTCG 7327251
 QY 504 AACCTAGCAGTCACTGGGCGC 526
 DB 7327250 CCGCGACACCTGTCTCATCGCC 7327228

RESULT 15

US-10-184-644-80/C
 Sequence 80, Application US/10184644
 Publication No. US20030044930A1
 GENERAL INFORMATION:
 APPLICANT: Baker, Kevin P.
 APPLICANT: Chen, Jian
 APPLICANT: Desnoyers, Luc
 APPLICANT: Goddard, Audrey
 APPLICANT: Godowski, Paul J.
 APPLICANT: Gurney, Austin L.
 APPLICANT: Pan, James
 APPLICANT: Smith, Victoria
 APPLICANT: Watanabe, Colin K.
 APPLICANT: Wood, William I.
 APPLICANT: Zhang, Zhenli
 TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
 FILE REFERENCE: P3430R1C227
 CURRENT APPLICATION NUMBER: US/10/184,644
 PRIOR FILING DATE: 2002-06-28
 PRIOR APPLICATION removed - See File Wrapper or Palm
 NUMBER OF SEQ ID NOS: 612
 SEQ ID NO 80
 LENGTH: 351
 TYPE: PRT
 ORGANISM: Homo Sapien
 US-10-184-644-80

Query Match 3.3%; Score 41.4; DB 14; Length 351;
 Best Local Similarity 13.4%; Pred. No. 0.0073;
 Matches 45; Conservative 103; Mismatches 187; Indels 0; Gaps 0;

QY 161 GAACCAATGATTAAGGTTATCGCGCTGACGAATCGCGGACGCTAATAAACCCTCCCTT 220
 DB 350 GAD..B.Y.G.YGMB.GGMD.MSGV.T.B.TAD.CYCG...GCH.TMSANNGAYCA 291
 QY 221 GCTGAGGTTGGGACATGACCCGATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 280
 DB 290 C.BCSYKH..G.H.S...RB.H.HAD.M..TNY.B.TYSSCBY..B.ATCH.M.M.TDC 231
 QY 281 GGCTACGAATGTTGGGACGCGCCACCTGTTTCCGCGCTGATTTCCGGGCGCTCACCGAA 340
 DB 230 B.Y...BTYGNAB.GSCSDGKG..BYKDA..TMYCT.NND.NK.TYSSSTCMYS.YKH 171
 QY 341 ATGCTGTCATCGCCCTGTTGTTGGGCGACGCGCCCTGAGCGCATCGGCTCACCA 400
 DB 170 B.S.S.TBCRT.NHSGCSW..C..SBCDBAYHGGM...CSGTYABT.TBCTSS.TRB. 111
 QY 401 TTGCTGTAAGTTCGCGCGACGATTTCTATGCGTTTCATTCGCGCGCATGTGTCAAA 460
 DB 110 TT.TAC.ABMC.B..SH.ASGMYTNY..R.G...T..TCYBSHKB..NYMK.B.Y. 51
 QY 461 AACCCATGCGCGCTTTCTATTGCTGCTGCT 495
 DB 50 ..CM.M.GA...S.SCBSS.MT.R.MNSTTTS.BT 16

Search completed: April 17, 2004, 21:01:41
 Job time: 642.884 secs

Matches 318; Conservative 0; Mismatches 296; Indels 3; Gaps 1;

```

Qy 222 CTGAGGTTGGGCGATGACCGGATGATGCGTTGCTCTTGGTATTCATACG 281
Db 53 CGGTGGTTGGGCGTGTATCCGCTGGGTTGGGCTTGGGCTTGGATGATGACAG 112
Qy 282 GCTACGATGATGAGGCGACCCCACTGTTTCCGCGCTGATTTCCGGGGCTCCACGAA 341
Db 113 GTTCCGCTGATGATGACCGGATTTCTCTTGGATGATGCGGTTGGATGAAAT 172
Qy 342 TGCTGATCATGCGCTGATGATGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 401
Db 173 TTCTGGCATGCGGATGATGACCGGATGATGCGGCGGCGGCGGCGGCGGCGGCGG 232
Qy 402 TGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 461
Db 233 TCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 292
Qy 462 ACCCATGCGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 521
Db 293 CCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 352
Qy 522 CGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 578
Db 353 CAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 412
Qy 579 ACTCTACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 638
Db 413 AAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 472
Qy 639 AATTAAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 698
Db 473 ATCTAAAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 532
Qy 699 GCCGACGAAAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 758
Db 533 TCATAAATTAAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 592
Qy 759 TTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 818
Db 593 GTTGTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 652
Qy 819 CCATCCGATCTCTTC 835
Db 653 TTCTCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 669

```

RESULT 12
US-09-738-626-1/c

Sequence 1, Application US/09738626
Publication No. US20020197605A1
GENERAL INFORMATION:
APPLICANT: NAKAGAWA, SATOSHI
APPLICANT: MIZOGUCHI, HIROSHI
APPLICANT: ANDO, SEIKO
APPLICANT: HAYASHI, MIKIRO
APPLICANT: OCHIAI, KEIKO
APPLICANT: YOKOI, HARUHIKO
APPLICANT: TATEISHI, NAOKO
APPLICANT: SENOH, AKIHIRO
APPLICANT: IKEDA, MASATO
APPLICANT: OZAKI, AKIO
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-125
CURRENT APPLICATION NUMBER: US/09/738, 626
PRIOR FILING DATE: 2000-12-18
PRIOR APPLICATION NUMBER: JP 99/377484
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: JP 00/159162
PRIOR FILING DATE: 2000-04-07
PRIOR APPLICATION NUMBER: JP 00/280988
PRIOR FILING DATE: 2000-08-03
NUMBER OF SEQ ID NOS: 7059

SOFTWARE: PatentIn ver. 3.0

SEQ ID NO 1
LENGTH: 3309400
TYPE: DNA
ORGANISM: Corynebacterium glutamicum
US-09-738-626-1

Query Match 10.0%; Score 127.4; DB 9; Length 3309400;
Best Local Similarity 51.5%; Pred. No. 1.4e-29;
Matches 318; Conservative 0; Mismatches 296; Indels 3; Gaps 1;

```

Qy 222 CTGAGGTTGGGCGATGACCGGATGATGCGTTGCTCTTGGTATTCATACG 281
Db 3289969 CGGTGGTTGGGCGTGTATCCGCTGGGTTGGGCTTGGGCTTGGATGATGACAG 3289910
Qy 282 GCTACGATGATGAGGCGACCCCACTGTTTCCGCGCTGATTTCCGGGGCTCCACGAA 341
Db 3289909 GTTCCGCTGATGATGACCGGATTTCTCTTGGATGATGCGGTTGGATGAAAT 3289850
Qy 342 TGCTGATCATGCGCTGATGATGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 401
Db 3289849 TTCTGGCATGCGGATGATGACCGGATGATGCGGCGGCGGCGGCGGCGGCGG 3289790
Qy 402 TGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 461
Db 3289789 TCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3289730
Qy 462 ACCCATGCGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 521
Db 3289729 CCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 3289670
Qy 522 CGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 578
Db 3289669 CAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 3289610
Qy 579 ACTCTACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 638
Db 3289609 AAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 638
Qy 639 AATTAAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 698
Db 3289549 ATCTAAAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 3289490
Qy 699 GCCGACGAAAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 758
Db 3289489 TCATAAATTAAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 3289430
Qy 759 TTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 818
Db 3289429 GTTGTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 3289370
Qy 819 CCATCCGATCTCTTC 835
Db 3289369 TTCTCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 3289353

```

RESULT 13

US-10-156-761-6067
Sequence 6067, Application US/10156761
Publication No. US20030119018A1
GENERAL INFORMATION:
APPLICANT: OMURA, SATOSHI
APPLICANT: ISHIKAWA, HARUO
APPLICANT: ISHIKAWA, UTN
APPLICANT: HORIKAWA, HIROSHI
APPLICANT: SHIBI, TADAYOSHI
APPLICANT: SAKAKI, YOSHIYUKI
APPLICANT: HATORI, MASAHIRA
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-262
CURRENT APPLICATION NUMBER: US/10/156,761
CURRENT FILING DATE: 2002-05-29
PRIOR APPLICATION NUMBER: JP 2001-204089

APPLICANT: SAKM, HERMANN
 APPLICANT: EGGELING, LOHAR
 APPLICANT: PFEFFERLE, WALTER
 TITLE OF INVENTION: NUCLEOTIDE SEQUENCE CODING FOR THE EXPORT OF
 TITLE OF INVENTION: BRANCHED CHAIN AMINO ACIDS, PROCESS FOR THE
 TITLE OF INVENTION: ISOLATION THEREOF AND USE THEREOF
 FILE REFERENCE: 7601/80525
 CURRENT APPLICATION NUMBER: US/10/608,504
 CURRENT FILING DATE: 2003-06-30
 PRIOR APPLICATION NUMBER: US 09/471,803
 PRIOR FILING DATE: 1999-12-23
 PRIOR APPLICATION NUMBER: DE 199 51 708.8
 PRIOR FILING DATE: 1999-10-27
 NUMBER OF SEQ ID NOS: 12
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO: 4
 LENGTH: 324
 TYPE: DNA
 ORGANISM: Corynebacterium glutamicum
 FEATURE:
 NAME/KEY: CDS
 LOCATION: (1)..(324)
 OTHER INFORMATION: brnE
 FEATURE:
 OTHER INFORMATION: ATCC14752
 US-10-608-504-4

Query Match 25.5%; Score 324; DB 15; Length 324;
 Best Local Similarity 100.0%; Pred. No. 4.3e-99;
 Matches 324; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 853 ATGACACTGATTTCTCCGTGATTTCTCCGTGTTGCGAGTATGCGATCAATTGCGG 912
 DB 1 ATGACACTGATTTCTCCGTGATTTCTCCGTGTTGCGAGTATGCGATCAATTGCGG 60
 QY 913 GCGCTCCGCGCGGCTTCCTTTTAACTCTTAAGCCCTTACGATCAATTGCGG 972
 DB 61 GCGCTCCGCGCGGCTTCCTTTTAACTCTTAAGCCCTTACGATCAATTGCGG 120
 QY 973 AAAATGGCATGTGATGCGACAGGATCTTGCAATTTGACGSCATCAAGTTTGGC 1032
 DB 121 AAAATGGCATGTGATGCGACAGGATCTTGCAATTTGACGSCATCAAGTTTGGC 180
 QY 1033 AGCAATGCGATGATCTGAAGACTTAACCTTTGCTTATGCGGCTTCCATTAAGTG 1092
 DB 181 AGCAATGCGATGATCTGAAGACTTAACCTTTGCTTATGCGGCTTCCATTAAGTG 240
 QY 1093 GTGGGCGCATCTTTGCGCGGTGAGACGACCTTGTTGAGCGTTGGCGCTGGACCATCGTT 1152
 DB 241 GTGGGCGCATCTTTGCGCGGTGAGACGACCTTGTTGAGCGTTGGCGCTGGACCATCGTT 300
 QY 1153 TTGTGGAAGTGTGATCTTTTC 1176
 DB 301 TTGTGGAAGTGTGATCTTTTC 324

RESULT 10
 US-10-134-640-1/c
 Sequence 1, Application US/10134640
 Publication No. US20030017555A1
 GENERAL INFORMATION:
 APPLICANT: Baehre, Brigitte
 APPLICANT: Kallnowski, Jörn
 APPLICANT: Puhler, Alfred
 APPLICANT: Mockel, Bettina
 APPLICANT: Pfeifferle, Walter
 TITLE OF INVENTION: Nucleotide Sequences Coding for the lrp Gene
 FILE REFERENCE: 990109 BT
 CURRENT APPLICATION NUMBER: US/10/134,640
 CURRENT FILING DATE: 2002-04-30
 NUMBER OF SEQ ID NOS: 9
 SOFTWARE: PatentIn version 3.1
 SEQ ID NO: 1

LENGTH: 715
 TYPE: DNA
 ORGANISM: Corynebacterium glutamicum
 FEATURE:
 NAME/KEY: CDS
 LOCATION: (151)..(612)
 OTHER INFORMATION: lrp-Gen
 FEATURE:
 NAME/KEY: -10 signal
 LOCATION: (88)..(93)
 OTHER INFORMATION:
 FEATURE:
 NAME/KEY: -35 signal
 LOCATION: (62)..(67)
 OTHER INFORMATION:
 US-10-134-640-1

Query Match 14.6%; Score 185; DB 12; Length 715;
 Best Local Similarity 100.0%; Pred. No. 1.5e-51;
 Matches 185; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCGGATCAATGGAATCTAGCTTCAATATATTGACAAATAGCCTAAGTTGAGTGGCAAC 60
 DB 185 GCGGATCAATGGAATCTAGCTTCAATATATTGACAAATAGCCTAAGTTGAGTGGCAAC 126
 QY 61 TGGCAAAACTACCCGCAATTTGTGTATGTTAGTGTGCAAAAAACGCAAGAGAT 120
 DB 125 TGGCAAAACTACCCGCAATTTGTGTATGTTAGTGTGCAAAAAACGCAAGAGAT 66
 QY 121 TCATTCAAGCTGGAAGTGTGCGCATCAAGGAGCCCTGGAACAGATGATTAAGTTA 180
 DB 65 TCATTCAAGCTGGAAGTGTGCGCATCAAGGAGCCCTGGAACAGATGATTAAGTTA 6
 QY 181 TCGGC 185
 DB 5 TCGGC 1

RESULT 11
 US-09-738-626-3408
 Sequence 3408, Application US/09738626
 Publication No. US20020197605A1
 GENERAL INFORMATION:
 APPLICANT: NAKAGAWA, SATOSHI
 APPLICANT: MIZOGUCHI, HIROSHI
 APPLICANT: ANDO, SEIKO
 APPLICANT: HAYASHI, MIKIRO
 APPLICANT: OCHIAI, KEIKO
 APPLICANT: YOKOI, HARUHIKO
 APPLICANT: TATEISHI, NAOKO
 APPLICANT: SENOH, AKIHIRO
 APPLICANT: IKEDA, MASATO
 APPLICANT: OZAKI, AKIO
 TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
 FILE REFERENCE: 249-125
 CURRENT APPLICATION NUMBER: US/09/738,626
 CURRENT FILING DATE: 2000-12-18
 PRIOR APPLICATION NUMBER: JP 99/377484
 PRIOR FILING DATE: 1999-12-16
 PRIOR APPLICATION NUMBER: JP 00/159162
 PRIOR FILING DATE: 2000-04-07
 PRIOR APPLICATION NUMBER: JP 00/280988
 PRIOR FILING DATE: 2000-08-03
 NUMBER OF SEQ ID NOS: 7059
 SOFTWARE: PatentIn ver. 3.0
 SEQ ID NO: 3408
 LENGTH: 711
 TYPE: DNA
 ORGANISM: Corynebacterium glutamicum
 US-09-738-626-3408

Query Match 10.0%; Score 127.4; DB 9; Length 711;
 Best Local Similarity 51.5%; Pred. No. 6.1e-32;

QY 9 AATGAGATCTAGCTCATATATATGCAACATAGCTGTAAGGTGGCAAACTGGCAAA 68

Db 376 AATGAACTAGCTTCATATATTGCAACATAGCTGTAAGGTGGCAAACTGGCAAA 317

QY 69 AAATCTACCCGGCACTTGTGTGATGATTTGATGTGCAAAAACGCAAGATTCATCTCA 128

Db 316 AAATCTACCCGGCACTTGTGTGATGATTTGATGTGCAAAAACGCAAGATTCATCTCA 257

QY 129 GCCTGAGAGTGTGCCATCCAAAGCAGCCCTGGAAACAGATGATAAAGTTATCGCGCT 188

Db 256 GCCTGAGAGTGTGCCATCCAAAGCAGCCCTGGAAACAGATGATAAAGTTATCGCGCT 197

QY 189 AGCAAAATCGGCAAGGCTCTAAAACTCCTTGTGTGCAAGTTTGGGCATGTACCAGATTG 248

Db 196 AGCAAAATCGGCAAGGCTCTAAAACTCCTTGTGTGAGTTTGGGCATGTACCAGATTG 137

QY 249 GATATTGCGTTGCTCTCTTGGTATTCAATACGCTACGATGATGTGGCAGCCCACTGT 308

Db 136 GATATTGCGTTGCTCTCTTGGTATTCAATACGCTACGATGATGTGGCAGCCCACTGT 77

QY 309 TTTTCGGGCTGATTTTTCGGGGGCTCCACGCAAAATGCGTGCATCGCCCTGTTGTGGGCG 368

Db 76 TTTTCGGGCTGATTTTTCGGGGGCTCCACGCAAAATGCGTGCATCGCCCTGTTGTGGGCG 17

RESULT 9
US-10-608-504-4
; Sequence 4, Application US/10608504
; Publication No. US2004001413A1
; GENERAL INFORMATION:
; APPLICANT: KENNERKNECHT, NICOLE

RESULT 5
US-10-608-504-2
Sequence 2, Application US/10608504
Publication No. US20040014123A1
GENERAL INFORMATION:
APPLICANT: KENNERKNECHT, NICOLE
APPLICANT: SAHM, HERMANN
APPLICANT: EGGELING, LOTMAR
APPLICANT: PFEFFERLE, WALTER
TITLE OF INVENTION: NUCLEOTIDE SEQUENCE CODING FOR THE EXPORT OF
TITLE OF INVENTION: BRANCHED CHAIN AMINO ACIDS, PROCESS FOR THE
FILE REFERENCE: 7601/80525
CURRENT APPLICATION NUMBER: US/10/608,504
CURRENT FILING DATE: 2003-06-30
PRIOR APPLICATION NUMBER: US 09/471,803
PRIOR FILING DATE: 1999-12-23
PRIOR APPLICATION NUMBER: DE 199 51 708.8
PRIOR FILING DATE: 1999-10-27
NUMBER OF SEQ ID NOS: 12
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO: 2
LENGTH: 753
TYPE: DNA
ORGANISM: Corynebacterium glutamicum
NAME/KEY: CDS
LOCATION: (1)..(753)
OTHER INFORMATION: brnp
FEATURE:
OTHER INFORMATION: ATCC14752
US-10-608-504-2

Query Match 59.2%; Score 753; DB 15; Length 753;
Best Local Similarity 100.0%; Pred. No. 6,5e-245; Indels 0; Gaps 0;
Matches 753; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

101 GTGCAAAAAGCGAAGATTGATTCAGAGCTGAGGTGTGCGCATCCAGGCGAGCTG 160
1 GTGCAAAAAGCGAAGATTGATTCAGAGCTGAGGTGTGCGCATCCAGGCGAGCTG 60
161 GAACGAGATGATTAAGGTATGAGGCTGATGAGGCTGATGAGGCTGATGAGGCTG 220
61 GAACGAGATGATTAAGGTATGAGGCTGATGAGGCTGATGAGGCTGATGAGGCTG 120
221 GGTGCAAGTTTGGGCAATGATACCGATGATGATGATGATGATGATGATGATGATG 280
121 GGTGCAAGTTTGGGCAATGATACCGATGATGATGATGATGATGATGATGATGATG 180
281 GGTGCAAGTTTGGGCAATGATACCGATGATGATGATGATGATGATGATGATGATG 340
181 GGTGCAAGTTTGGGCAATGATACCGATGATGATGATGATGATGATGATGATGATG 240
341 ATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 400
241 ATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 300
401 TTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 460
301 TTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 360
461 AACCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 520
361 AACCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 420
521 GCGGCGAGGCGGCGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATG 580
421 GCGGCGAGGCGGCGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATG 480
581 TTCTAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 640
481 TTCTAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 540

QY 641 ATTAAGGCGCTGAGTGGAGCCCTTGTCTCTCTTGTGTCAGCTGATCTGATCTGTC 700
DB 541 ATTAAGGCGCTGAGTGGAGCCCTTGTCTCTCTTGTGTCAGCTGATCTGATCTGTC 600
QY 701 CGAAGCAAAAAGCAGATCCCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 760
DB 601 CGAAGCAAAAAGCAGATCCCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
QY 761 GTGCTAATTCAGATCAGGCGCCCTATTTGGGCGCTGATGATGATGATGATGATG 820
DB 661 GTGCTAATTCAGATCAGGCGCCCTATTTGGGCGCTGATGATGATGATGATGATG 720
QY 821 ATCCGCTACTTCTTCTTGGGAAAAGCTGCTAAA 853
DB 721 ATCCGCTACTTCTTCTTGGGAAAAGCTGCTAAA 753

RESULT 6
US-10-134-640-7/c
Sequence 7, Application US/10134640
Publication No. US20030017555A1
GENERAL INFORMATION:
APPLICANT: Bache, Brigitte
APPLICANT: Kalinowski, Jörn
APPLICANT: Punier, Alfred
APPLICANT: Mockel, Bettina
APPLICANT: Pfefferle, Walter
TITLE OF INVENTION: Nucleotide Sequences Coding for the lrp Gene
FILE REFERENCE: 990109 BT
CURRENT APPLICATION NUMBER: US/10/134,640
CURRENT FILING DATE: 2002-04-30
NUMBER OF SEQ ID NOS: 9
SOFTWARE: PatentIn version 3.1
SEQ ID NO: 7
LENGTH: 397
TYPE: DNA
ORGANISM: Corynebacterium glutamicum
NAME/KEY: misc_feature
LOCATION: (1)..(397)
OTHER INFORMATION: lrp part 1
US-10-134-640-7

Query Match 28.8%; Score 365.6; DB 12; Length 397;
Best Local Similarity 98.9%; Pred. No. 3,3e-113; Indels 0; Gaps 0;
Matches 368; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 9 AATGAAATCTAGCTTCAATATATGACAAATAGCTTGAAGTGGCCAAACTGGCAACA 68
DB 376 AATGAAATCTAGCTTCAATATATGACAAATAGCTTGAAGTGGCCAAACTGGCAACA 317
QY 69 AAATACCCGCGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 128
DB 316 AAATACCCGCGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 257
QY 129 GCTGAGATGCTGCGCATTCAGAGGAGCCCTGGAACCAAGATGATTAAGTTAGCGGCT 188
DB 256 GCTGAGATGCTGCGCATTCAGAGGAGCCCTGGAACCAAGATGATTAAGTTAGCGGCT 197
QY 189 ACGAAATGCGCAAGATCTTAAACCTCCCTGCTGAGAGTTGGGAGATGATCCGATG 248
DB 196 ACGAAATGCGCAAGATCTTAAACCTCCCTGCTGAGAGTTGGGAGATGATCCGATG 137
QY 249 GTATGAGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 308
DB 136 GTATGAGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 77
QY 309 TTTCGAGCTGATTTTGGCGGCTGACCGAAATGCTGATGATGATGATGATGATGATG 368
DB 76 TTTCGAGCTGATTTTGGCGGCTGACCGAAATGCTGATGATGATGATGATGATGATG 17
QY 369 CAGCGCCCTGG 380

```

Db 277209 TTGGATTTTGGGCTTATGAGAGAGGCTTACGACATCTGCGCCAGGCCCCGAGGCTG 277268
QY 541 GTGCGCGTGGGCACTTATCTCAATGCAAAATAGCGTTTCACTCTTACCTGAGTTCGGCGG 600
Db 277269 GTGCGCGTGGGCACTTATCTCAATGCAAAATAGCGTTTCACTCTTACCTGAGTTCGGCGG 277328
QY 601 TCTCACCGGAGTGGGAGTTCGAGATGATTCCTTTAAATTAAAGGCGCTCGAGTTGCG 660
Db 277329 TCTCACCGGAGTGGGAGTTCGAGATGATTCCTTTAAATTAAAGGCGCTCGAGTTGCG 277388
QY 661 CTTTGTCTCTCTTTTGTGACGCTGACTTTGAGATTCCTGCGGAGGAGGAGGAGATCCC 720
Db 277389 CTTTGTCTCTCTTTTGTGACGCTGACTTTGAGATTCCTGCGGAGGAGGAGATCCC 277448
QY 721 TTCTCTGCTGCTGCGAGGTTTGAAGTTTCAACATGCTCTTGTGTATTTTCAAGGTGAGG 780
Db 277449 TTCTCTGCTGCTGCGAGGTTTGAAGTTTCAACATGCTCTTGTGTATTTTCAAGGTGAGG 277508
QY 781 CCTATTGCGGCGCTGCTGATCTTCTTGGGCTGTTGACCAATCCGCTACTTCTTCTGGG 840
Db 277509 CCTATTGCGGCGCTGCTGATCTTCTTGGGCTGTTGACCAATCCGCTACTTCTTCTGGG 277568
QY 841 AAAGGCTGCTAAATGACAACTGATTTCTCTGTATTTCTCTTGTGTTGTCGAGTATGTCGA 900
Db 277569 AAAGGCTGCTAAATGACAACTGATTTCTCTGTATTTCTCTTGTGTTGTCGAGTATGTCGA 277628
QY 901 GTGATTAATTTTGGGCTGCGGCGGCTGCTGCTTCTTAACTTAAAGCCCTACGTAATCA 960
Db 277629 GTGATTAATTTTGGGCTGCGGCGGCTGCTGCTTCTTAACTTAAAGCCCTACGTAATCA 277688
QY 961 CAATTGTGTGGGCAAAATGCGGATGTGAGATGCGAGCAAGAACTCTTGCCATTTTGACCGGA 1020
Db 277689 CAATTGTGTGGGCAAAATGCGGATGTGAGATGCGAGCAAGAACTCTTGCCATTTTGACCGGA 277748
QY 1021 TCAACGTTTGGCGAGATGCGATGAGATCTGAAGACCTTAACTTAAAGCCCTACGTAATCA 1080
Db 277749 TCAACGTTTGGCGAGATGCGATGAGATCTGAAGACCTTAACTTAAAGCCCTACGTAATCA 277808
QY 1081 GCGATTAACAGTGTGCGGCACTTCTTGGGCGTGAAGCACTTGTGAGCGTTGAGCGCT 1140
Db 277809 GCGATTAACAGTGTGCGGCACTTCTTGGGCGTGAAGCACTTGTGAGCGTTGAGCGCT 277868
QY 1141 GGCACCACTGTTTTTGTGAGACTGTGGAATCTTTCTTAAACGTAATTAACAAAAT 1200
Db 277869 GGCACCACTGTTTTTGTGAGACTGTGGAATCTTTCTTAAACGTAATTAACAAAAT 277928
QY 1201 CCGCATGCGCTCAATTTGAAGGGGATGCGGATTTTGAAGAACTTGAAGAAAGGCTTAAG 1260
Db 277929 CCGCATGCGCTCAATTTGAAGGGGATGCGGATTTTGAAGAACTTGAAGAAAGGCTTAAG 277988
QY 1261 CAGACAGCGCT 1271
Db 277989 CAGACAGCGCT 277999

```

```

RESULT 4
US-09-738-626-289
; Sequence 289, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAOKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125

```

```

; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 289
; LENGTH: 753
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-289

Query Match 59.2%; Score 753; DB 9; Length 753;
Best Local Similarity 100.0%; Pred. No. 6.5e-245;
Matches 753; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 101 GTGCAAAAACGCAAGATTCATTCAAGCTTGAAGTGTGCTGATCCAAAGGAGGCTTG 160
Db 1 GTGCAAAAACGCAAGATTCATTCAAGCTTGAAGTGTGCTGATCCAAAGGAGGCTTG 60
QY 161 GAACCAATGATTAAGGTTATCGGCGCTACGAATGCGCAAGGCTTAAAACTCCCTT 220
Db 61 GAACCAATGATTAAGGTTATCGGCGCTACGAATGCGCAAGGCTTAAAACTCCCTT 120
QY 221 GCTGCAAGTTTGGGCAATGACCCGATTTGATTTGCTTGTGCTTCTTGTATTAATAC 280
Db 121 GCTGCAAGTTTGGGCAATGACCCGATTTGATTTGCTTGTGCTTCTTGTATTAATAC 160
QY 281 GGTACGAATGTGGGAGGAGCCCACTGTTTCCGCGCTGATTTTCCGCGGCTCCACCGAA 340
Db 181 GGTACGAATGTGGGAGGAGCCCACTGTTTCCGCGCTGATTTTCCGCGGCTCCACCGAA 240
QY 341 ATGCTGTGATCGCCCTGTTGTGTGGGCGAGCGCCCTGAGCGCCATCGGCTCACCA 400
Db 241 ATGCTGTGATCGCCCTGTTGTGTGGGCGAGCGCCCTGAGCGCCATCGGCTCACCA 300
QY 401 TTGCTGTGAACTTCGCGCAAGTATTTATGCGTTTCAATCCGCTGATGAGTCAAA 460
Db 301 TTGCTGTGAACTTCGCGCAAGTATTTATGCGTTTCAATCCGCTGATGAGTCAAA 360
QY 461 AACCCATTTGCGCTTCTATTTCCGTTTTCGCGCTTATGACGAAGCTTACGAGTACT 520
Db 361 AACCCATTTGCGCTTCTATTTCCGTTTTCGCGCTTATGACGAAGCTTACGAGTACT 420
QY 521 GCGGCGAGGCGCGAGGCTGTGCTGCGGCTGCGGATTTTCAATGCAATAGCGTTTAC 580
Db 421 GCGGCGAGGCGCGAGGCTGTGCTGCGGCTGCGGATTTTCAATGCAATAGCGTTTAC 480
QY 581 TCTTACTGAGTATTCGCGGCTCTCACCGAGTGTGCGAGTGTGATTCCTTTGAA 640
Db 481 TCTTACTGAGTATTCGCGGCTCTCACCGAGTGTGCGAGTGTGATTCCTTTGAA 540
QY 641 ATTAAGGCGCTGAGTTGCGCTTGTCTCTCTTGTCAAGCTGATCTTGTGATTCCTG 700
Db 541 ATTAAGGCGCTGAGTTGCGCTTGTCTCTCTTGTCAAGCTGATCTTGTGATTCCTG 600
QY 701 CGAAGCAAAAGCGATCCCTCTCTGCTGCTGCTGAGGTTTGAAGTTCACCATGCTCTT 760
Db 601 CGAAGCAAAAGCGATCCCTCTCTGCTGCTGCTGAGGTTTGAAGTTCACCATGCTCTT 660
QY 761 GTGTAATTCAGGTCAGGCTTATTTGCGCGCTGCTGATCTTGTGAGTCTGTTGAC 820
Db 661 GTGTAATTCAGGTCAGGCTTATTTGCGCGCTGCTGATCTTGTGAGTCTGTTGAC 720
QY 821 ATCCGTAATCTTCTTGGGAAAGGCTGCTAAA 853
Db 721 ATCCGTAATCTTCTTGGGAAAGGCTGCTAAA 753

```

Db 181 TCGGGGCTACGAAATCGCGCAAGGCTTAAACCTCCCTTGTGAGGTTGGGATGTA 240
QY 241 CCCGATGATATGCGTTGATCTCTGATTAATCAATACGAGCTACGATGATGGGAGC 300
Db 241 CCCGATGATATGCGTTGATCTCTGATTAATCAATACGAGCTACGATGATGGGAGC 300
QY 301 CCCGATGATATGCGTTGATCTCTGATTAATCAATACGAGCTACGATGATGGGAGC 360
Db 301 CCCGATGATATGCGTTGATCTCTGATTAATCAATACGAGCTACGATGATGGGAGC 360
QY 361 TGTGGGCGAGGCGCCCTGAGGCGCATGCGGCTCACACATGCTGTGAATCTTCGCGCA 420
Db 361 TGTGGGCGAGGCGCCCTGAGGCGCATGCGGCTCACACATGCTGTGAATCTTCGCGCA 420
QY 421 CGATTTCTATGCGTTTCAATCCCGCTGATGATGATCAAAACCCCATTTGCCGTTCTA 480
Db 421 CGATTTCTATGCGTTTCAATCCCGCTGATGATGATCAAAACCCCATTTGCCGTTCTA 480
QY 481 TTGGGTTTTCGCGCTTATGAGAGACCTACGAGCTACGAGGCGAGGCGCGAGCTG 540
Db 481 TTGGGTTTTCGCGCTTATGAGAGACCTACGAGCTACGAGGCGAGGCGCGAGCTG 540
QY 541 GTGGGCGTGGCACTTATCTCAATGCAATAGCGTTTCACTCTACTGGGATTCGCGCG 600
Db 541 GTGGGCGTGGCACTTATCTCAATGCAATAGCGTTTCACTCTACTGGGATTCGCGCG 600
QY 601 TCTCACCAGATGCGCATGCGAGATGATTCCTTTTGAATTAAGGCGCTGAGTGGC 660
Db 601 TCTCACCAGATGCGCATGCGAGATGATTCCTTTTGAATTAAGGCGCTGAGTGGC 660
QY 661 CCTTGTCTCTCTTTGTGACGCTGATGATTCCTGCGCAAGAAACGAAACGATGCC 720
Db 661 CCTTGTCTCTCTTTGTGACGCTGATGATTCCTGCGCAAGAAACGAAACGATGCC 720
QY 721 TTCTGTCTGTGCGAGGTTTGAAGCTTCAACATGCTCTTGTGATTAATTCAGTCCAGC 780
Db 721 TTCTGTCTGTGCGAGGTTTGAAGCTTCAACATGCTCTTGTGATTAATTCAGTCCAGC 780
QY 781 CCTATTGCGGCGCTGATCTCTTGGGCTGATGATTCCTGCGCAAGAAACGAAACGATGCC 840
Db 781 CCTATTGCGGCGCTGATCTCTTGGGCTGATGATTCCTGCGCAAGAAACGAAACGATGCC 840
QY 841 AAAGGCTGTAATGACATGATTCCTGATTCCTCTGATTCCTCTGATTCCTCTGATTCCT 900
Db 841 AAAGGCTGTAATGACATGATTCCTGATTCCTCTGATTCCTCTGATTCCTCTGATTCCT 900
QY 901 GTGATTAATCTTTGGGCTGCGGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTG 960
Db 901 GTGATTAATCTTTGGGCTGCGGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTG 960
QY 961 CAATTTGTGGGCAAAATGCGATGATGATGCGAGCAAGAACTCTTGGCACTTTTGAACGCA 1020
Db 961 CAATTTGTGGGCAAAATGCGATGATGATGCGAGCAAGAACTCTTGGCACTTTTGAACGCA 1020
QY 1021 TCAACGTTTTCGAGCAATGCGATGATGATGCGAGCAAGAACTCTTGGCACTTTTGAACGCA 1080
Db 1021 TCAACGTTTTCGAGCAATGCGATGATGATGCGAGCAAGAACTCTTGGCACTTTTGAACGCA 1080
QY 1081 GCGATTAATGATGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTG 1140
Db 1081 GCGATTAATGATGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTG 1140
QY 1141 GCGACATGCTTTTGTGAGCTGATGATTCCTTTTGAATTAAGGCGCTGAGTGA 1200
Db 1141 GCGACATGCTTTTGTGAGCTGATGATTCCTTTTGAATTAAGGCGCTGAGTGA 1200
QY 1201 CCGGATGCTGATTTGAAGGCGATGCGGATTTTGAAGGCGATTTTGAAGGCGATTTGA 1260
Db 1201 CCGGATGCTGATTTGAAGGCGATGCGGATTTTGAAGGCGATTTTGAAGGCGATTTGA 1260
QY 1261 CAGACAGCGCT 1271

Db 1261 CAGACAGCGCT 1271
RESULT 3
US-09-738-626-1
Sequence 1, Application US/09738626
Publication No. US20020197605A1
GENERAL INFORMATION:
APPLICANT: NAKAGAWA, SATOSHI
APPLICANT: MIZOGUCHI, HIROSHI
APPLICANT: ANDO, SEIKO
APPLICANT: HAYASHI, MIKIRO
APPLICANT: OCHIAI, KEIKO
APPLICANT: YOKOI, HARUHIKO
APPLICANT: TATEISHI, MAKOTO
APPLICANT: ISENOH, AKIHITO
APPLICANT: IKEDA, MASATO
APPLICANT: OZAKI, AKIO
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-125
CURRENT APPLICATION NUMBER: US/09/738,626
CURRENT FILING DATE: 2000-12-18
PRIOR APPLICATION NUMBER: JP 99/377484
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: JP 00/159162
PRIOR FILING DATE: 2000-04-07
PRIOR APPLICATION NUMBER: JP 00/280988
NUMBER OF SEQ ID NOS: 7059
SOFTWARE: PatentIn ver. 3.0
SEQ ID NO 1
LENGTH: 3309400
TYPE: DNA
ORGANISM: Corynebacterium glutamicum
US-09-738-626-1
Query Match 100.0%; Score 1271; DB 9; Length 3309400;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1271; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GCGGATCAATGGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 60
Db 276729 GCGGATCAATGGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 276788
QY 61 TGGCAACAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 120
Db 276789 TGGCAACAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 276848
QY 121 TCAATTAAGCTGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 180
Db 276849 TCAATTAAGCTGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 276908
QY 181 TCGGCTACGAAATGCGGAGCTGTAATAAACCCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
Db 276909 TCGGCTACGAAATGCGGAGCTGTAATAAACCCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 276968
QY 241 CCCGATGATATGCGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300
Db 276969 CCCGATGATATGCGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 277028
QY 301 CCCGATGATATGCGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 360
Db 277029 CCCGATGATATGCGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 277088
QY 361 TGTGGGCGAGGCGCCCTGAGGCGCATGCGGCTCACACATGCTGTGAATCTTCGCGCA 420
Db 277089 TGTGGGCGAGGCGCCCTGAGGCGCATGCGGCTCACACATGCTGTGAATCTTCGCGCA 277148
QY 421 CGATTTCTATGCGTTTCAATCCCGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
Db 277149 CGATTTCTATGCGTTTCAATCCCGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 277208
QY 481 TTGGGTTTTCGCGCTTATGAGAGACCTACGAGCTACGAGGCGAGGCGCGAGCTG 540

Best Local Similarity 100.0%; Pred. No. 0;
Matches 1271; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy 1 GGGCGATCAATGGAGATCTAGCTTCAATATTTGCAAAATAGCTTGAAGTGGGCAAC 60
Db 1 GGGCGATCAATGGAGATCTAGCTTCAATATTTGCAAAATAGCTTGAAGTGGGCAAC 60
Qy 61 TGGCAACAAATACCCGGCAATGTGTGATGTGTGATGTGTGATGTGTGATGTGTGAT 120
Db 61 TGGCAACAAATACCCGGCAATGTGTGATGTGTGATGTGTGATGTGTGATGTGTGAT 120
Qy 121 TCATTCAAGCTTGAGAGTGTGATGTGTGATGTGTGATGTGTGATGTGTGATGTGTGAT 180
Db 121 TCATTCAAGCTTGAGAGTGTGATGTGTGATGTGTGATGTGTGATGTGTGATGTGTGAT 180
Qy 181 TGGGCGCTACGAATGGCGCAAGTGTGATGTGTGATGTGTGATGTGTGATGTGTGAT 240
Db 181 TGGGCGCTACGAATGGCGCAAGTGTGATGTGTGATGTGTGATGTGTGATGTGTGAT 240
Qy 241 CCCGATTGATTCGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 300
Db 241 CCCGATTGATTCGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 300
Qy 301 CCCGATTGATTCGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 360
Db 301 CCCGATTGATTCGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 360
Qy 361 TGTGGCGAGAGCCCTTGGCGCATCGCGCTACCAACATTCCTGTGTGTGTGTGTGTGT 420
Db 361 TGTGGCGAGAGCCCTTGGCGCATCGCGCTACCAACATTCCTGTGTGTGTGTGTGTGT 420
Qy 421 CGATTCTATGCTTTTCAATTCCTGCTGATGTGTGTGTGTGTGTGTGTGTGTGTGTGT 480
Db 421 CGATTCTATGCTTTTCAATTCCTGCTGATGTGTGTGTGTGTGTGTGTGTGTGTGTGT 480
Qy 481 TTTGGTTTTCGCTTATCGACGAAAGCTACGAGTCACTGTGGCGGCGCGGCGAGCTG 540
Db 481 TTTGGTTTTCGCTTATCGACGAAAGCTACGAGTCACTGTGGCGGCGCGGCGAGCTG 540
Qy 541 GTGCGGCTGCGCATTAATCAATGCAAAATAGGCTTCTTCTTGTGTGTGTGTGTGTGT 600
Db 541 GTGCGGCTGCGCATTAATCAATGCAAAATAGGCTTCTTCTTGTGTGTGTGTGTGTGT 600
Qy 601 TCTCACCGAGTGGCGATCGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 660
Db 601 TCTCACCGAGTGGCGATCGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 660
Qy 661 CTTTGTCTCTCTTTGTGACGCTGACTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 720
Db 661 CTTTGTCTCTCTTTGTGACGCTGACTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 720
Qy 721 TTTCTGCTGCTGCGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 780
Db 721 TTTCTGCTGCTGCGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 780
Qy 781 CCTATTTCGCGCGCTGCTGATCTTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 840
Db 781 CCTATTTCGCGCGCTGCTGATCTTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 840
Qy 841 AAAGGCTGCTAAATGCAATGATTTCTGCTGATTTCTGCTGATTTCTGCTGATTTCTG 900
Db 841 AAAGGCTGCTAAATGCAATGATTTCTGCTGATTTCTGCTGATTTCTGCTGATTTCTG 900
Qy 901 GTCAATTCCTTGTGCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960
Db 901 GTCAATTCCTTGTGCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960
Qy 961 CAATTGTGGGCAAAATGGCGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1020
Db 961 CAATTGTGGGCAAAATGGCGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1020
Qy 1021 TCAACGTTTCGACGAAATGCGATGATCTGAAAGCTTAACTTTGTGTGTGTGTGTGT 1080
Db 1021 TCAACGTTTCGACGAAATGCGATGATCTGAAAGCTTAACTTTGTGTGTGTGTGTGT 1080

```

```

Db 1021 TCAACGTTTCGACGAAATGCGATGATCTGAAAGCTTAACTTTGTGTGTGTGTGT 1080
Qy 1081 GCGATTACAGATGTGGCGCATCTTCTTGGCGGTGACGACCTTGTGTGAGCGTGGCGCT 1140
Db 1081 GCGATTACAGATGTGGCGCATCTTCTTGGCGGTGACGACCTTGTGTGAGCGTGGCGCT 1140
Qy 1141 GCGACCAATGCTTTTGTGTGATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1200
Db 1141 GCGACCAATGCTTTTGTGTGATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1200
Qy 1201 CCGATGCGCTCAATTTTGAAGGGGATCGGATTTTGAAGAACTGAGAAAGGCTTAA 1260
Db 1201 CCGATGCGCTCAATTTTGAAGGGGATCGGATTTTGAAGAACTGAGAAAGGCTTAA 1260
Qy 1261 CAGACAGCGCT 1271
Db 1261 CAGACAGCGCT 1271

```

RESULT 2

```

US-10-608-504-6
/ Sequence 6, Application US/10608504
/ Publication No. US20040014123A1
/ GENERAL INFORMATION:
/ APPLICANT: KENNERNECHT, NICOLE
/ APPLICANT: SAHM, HERMANN
/ APPLICANT: EGGELING, LOTHAR
/ APPLICANT: PFERLE, WALTER
/ TITLE OF INVENTION: NUCLEOTIDE SEQUENCE CODING FOR THE EXPORT OF
/ TITLE OF INVENTION: BRANCHED CHAIN AMINO ACIDS, PROCESS FOR THE
/ TITLE OF INVENTION: ISOLATION THEREOF AND USE THEREOF
/ FILE REFERENCE: 7601/80525
/ CURRENT APPLICATION NUMBER: US/10/608,504
/ PRIOR APPLICATION NUMBER: US 09/471,803
/ PRIOR FILING DATE: 1999-12-23
/ PRIOR APPLICATION NUMBER: DE 199 51 708.8
/ PRIOR FILING DATE: 1999-10-27
/ NUMBER OF SEQ ID NOS: 12
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 6
/ LENGTH: 1271
/ TYPE: DNA
/ ORGANISM: Corynebacterium glutamicum
/ FEATURE:
/ NAME/KEY: gene
/ LOCATION: (101)..(853)
/ OTHER INFORMATION: brnF
/ FEATURE:
/ NAME/KEY: gene
/ LOCATION: (853)..(1176)
/ OTHER INFORMATION: brnE
/ FEATURE:
/ OTHER INFORMATION: ATCC33032
US-10-608-504-6

```

```

Query Match 100.0%; Score 1271; DB 15; Length 1271;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1271; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GGGCGATCAATGGAGATCTAGCTTCAATATTTGCAAAATAGCTTGAAGTGGGCAAC 60
Db 1 GGGCGATCAATGGAGATCTAGCTTCAATATTTGCAAAATAGCTTGAAGTGGGCAAC 60
Qy 61 TGGCAACAAATACCCGGCAATGTGTGATGTGTGATGTGTGATGTGTGATGTGTGAT 120
Db 61 TGGCAACAAATACCCGGCAATGTGTGATGTGTGATGTGTGATGTGTGATGTGTGAT 120
Qy 121 TCATTCAAGCTTGAGAGTGTGATGTGTGATGTGTGATGTGTGATGTGTGATGTGTGAT 180
Db 121 TCATTCAAGCTTGAGAGTGTGATGTGTGATGTGTGATGTGTGATGTGTGATGTGTGAT 180
Qy 181 TGGGCGCTACGAATGGCGCAAGTGTGTGATGTGTGATGTGTGATGTGTGATGTGTGAT 240
Db 181 TGGGCGCTACGAATGGCGCAAGTGTGTGATGTGTGATGTGTGATGTGTGATGTGTGAT 240

```

Tue Apr 20 06:47:23 2004

us-10-608-504-6.rnpb

Page 1

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 17, 2004, 06:48:18 ; Search time 617.884 Seconds
(without alignments)
9204.305 Million cell updates/sec

Title: US-10-608-504-6
Perfect score: 1271
Sequence: 1 gcgcagcatcagatcagctag.....agcttaacagacagcgcct 1271

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 2890132 seqs, 2237290429 residues

Total number of hits satisfying chosen parameters: 5780264

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:*

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
17: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
18: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1271	100.0	1271	15	US-10-608-504-1
2	1271	100.0	1271	15	US-10-608-504-6
3	1271	100.0	3309400	9	US-09-738-626-1
4	753	59.2	753	9	US-09-738-626-289
5	753	59.2	753	15	US-10-608-504-2
6	365.6	28.8	397	12	US-10-134-640-7
7	365.6	28.8	778	12	US-10-134-640-9
8	324	25.5	324	9	US-09-738-626-290
9	324	25.5	324	15	US-10-608-504-4
10	185	14.6	715	12	US-10-134-640-1
11	127.4	10.0	711	9	US-09-738-626-3408
12	127.4	10.0	3309400	9	US-09-738-626-1
13	43	3.4	594	14	US-10-156-761-6067
14	43	3.4	9025608	14	US-10-156-761-1
15	41.4	3.3	351	14	US-10-184-644-80

C 16	41.4	3.3	351	14	US-10-184-634-80	Sequence 80, Appl
C 17	40	3.1	671	14	US-10-184-644-346	Sequence 346, App
C 18	40	3.1	671	14	US-10-184-634-346	Sequence 346, App
C 19	39.8	3.1	596	14	US-10-184-644-310	Sequence 310, App
C 20	39.8	3.1	596	14	US-10-184-634-310	Sequence 310, App
C 21	39.8	3.1	596	14	US-10-063-688-100	Sequence 100, App
C 22	38.4	3.0	594	12	US-10-142-828-10	Sequence 10, Appl
C 23	38.4	3.0	594	14	US-10-123-155-10	Sequence 10, Appl
C 24	38.4	3.0	594	14	US-10-146-731-10	Sequence 10, Appl
C 25	38.4	3.0	594	14	US-10-140-472-10	Sequence 10, Appl
C 26	38.4	3.0	594	14	US-10-141-761-10	Sequence 10, Appl
C 27	38.4	3.0	594	14	US-10-142-888-10	Sequence 10, Appl
C 28	38.4	3.0	594	14	US-10-158-790-10	Sequence 10, Appl
C 29	38.4	3.0	594	15	US-10-137-871-10	Sequence 10, Appl
C 30	38.4	3.0	594	15	US-10-140-923-10	Sequence 10, Appl
C 31	38.4	3.0	594	15	US-10-141-756-10	Sequence 10, Appl
C 32	38.4	3.0	594	15	US-10-141-759-10	Sequence 10, Appl
C 33	38.4	3.0	594	15	US-10-140-805-10	Sequence 10, Appl
C 34	38.4	3.0	594	15	US-10-140-864-10	Sequence 10, Appl
C 35	38.2	3.0	946	9	US-09-778-320-211	Sequence 211, App
C 36	38.2	3.0	946	9	US-09-910-689-211	Sequence 211, App
C 37	38.2	3.0	946	13	US-10-010-742-211	Sequence 48, Appl
C 38	37.8	3.0	457	14	US-10-184-644-48	Sequence 48, Appl
C 39	37.8	3.0	457	14	US-10-184-634-48	Sequence 12, Appl
C 40	37.8	3.0	457	14	US-10-063-685-12	Sequence 326, App
C 41	37.4	2.9	557	14	US-10-184-644-326	Sequence 326, App
C 42	37.4	2.9	557	14	US-10-184-634-326	Sequence 5719, App
C 43	37	2.9	186	14	US-10-156-761-5719	Sequence 1, Appl
C 44	37	2.9	9025608	14	US-10-156-761-1	Sequence 30075, A
C 45	36.8	2.9	1230	12	US-10-282-122A-30075	

ALIGNMENTS

RESULT 1
US-10-608-504-1
Sequence 1, Application US/10608504
Publication No. US20040014123A1
GENERAL INFORMATION:
APPLICANT: KENNERKNECHT, NICOLE
APPLICANT: SAMM, HERMANN
APPLICANT: EGELING, WALTER
APPLICANT: PEEFERLE, WALTER
TITLE OF INVENTION: NUCLEOTIDE SEQUENCE CODING FOR THE EXPORT OF
TITLE OF INVENTION: BRANCHED CHAIN AMINO ACIDS, PROCESS FOR THE
TITLE OF INVENTION: ISOLATION THEREOF AND USE THEREOF
FILE REFERENCE: 7601/80525
CURRENT APPLICATION NUMBER: US/10/608,504
CURRENT FILING DATE: 2003-06-30
PRIOR APPLICATION NUMBER: US 09/471,803
PRIOR FILING DATE: 1999-12-23
PRIOR APPLICATION NUMBER: DE 199 51 708.8
PRIOR FILING DATE: 1999-10-27
NUMBER OF SEQ ID NOS: 12
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1
LENGTH: 1271
TYPE: DNA
ORGANISM: Corynebacterium glutamicum
FEATURE:
NAME/KEY: gene
LOCATION: (101)..(853)
OTHER INFORMATION: brnf
FEATURE:
NAME/KEY: gene
LOCATION: (853)..(1176)
OTHER INFORMATION: brne
FEATURE:
OTHER INFORMATION: ATCC14752
US-10-608-504-1
Query Match 100.0%; Score 1271; DB 15; Length 1271;

Tue Apr 20 06:47:23 2004

us-10-608-504-6.rn1

Page 8

TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 751 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-09-030-607-12

Query Match 2.9%; Score 36.4; DB 3; Length 751;
Best Local Similarity 59.8%; Pred. No. 0.099;
Matches 58; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

QY	327	CGGCGCTCCACCGAAATGCTGTCATGCGCCCTGTTGTGGCGCAGGCGCCCTGGGCGCCA	386
DB	410	CGTCTTCTTCATCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCT	351
QY	387	TCGCGCTCACACACATGCTGCTGAACCTCCGCCACGT	423
DB	350	TGCTGTACACCAATGCTGAGCACTTCCTGACGT	314

Search completed: April 19, 2004, 18:38:21
Job time : 113.018 secs

Db 593 ACGGATGAGCTTGGCGCCG 612

RESULT 9

US-09-252-991A-11418
Sequence 11418, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 11418
LENGTH: 1086
TYPE: DNA
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-11418

Query Match
Best Local Similarity 3.1%; Score 39.2; DB 4; Length 1086;
Matches 77; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

Qy 309 TTTCGGGCTGATTTTCGGGGCTCCACCAATGCTGTCATCGCCCTCGTTGGGCG 368
Db 222 TGTCGCCATCGTCTTCGCCGGCCGCGCAGTTGTCGATCGCATGCTCAAGGCG 281

Qy 369 CACGGCCCTGGGGCGCATGCGCTCACCAATGCTGTCGATGCTCCGACGATTTCT 428
Db 282 GTGCCGGGTTCTTTGATATCTTACCAACCTCTCTGACCTCGACGACCTGCTCT 341

Qy 429 ATGCGTTTCATTCGCGCTG 448
Db 342 ACGGATGAGCTTGGCGCCG 361

RESULT 10
US-09-621-976-8976/c
Sequence 8976, Application US/09621976
Patent No. 6639063
GENERAL INFORMATION:
APPLICANT: Dumas Milne Edwards, J.B.
APPLICANT: Jobert, S.
APPLICANT: Giordano, J.Y.
TITLE OF INVENTION: ESTs and Encoded Human Proteins.
FILE REFERENCE: GENSET.054PR2
CURRENT APPLICATION NUMBER: US/09/621,976
CURRENT FILING DATE: 2000-07-21
NUMBER OF SEQ ID NOS: 19335
SOFTWARE: Patent.pm
SEQ ID NO 8976
LENGTH: 399
TYPE: DNA
ORGANISM: Homo sapiens
US-09-621-976-8976

Query Match
Best Local Similarity 3.1%; Score 39; DB 4; Length 399;
Matches 30; Conservative 112; Mismatches 97; Indels 0; Gaps 0;

Qy 609 GAGTGGGATGCGAGTGAATTCCTTTGAATAAGGCTCGAGTCCCTTTGCT 668
Db 242 SAMRRKKMTGKRSWSGSRSTGYAMMYKSWCTSRKMYKXKRRKXCTSTKT 183
Qy 669 CTCCTTTGACGCTGACTTGAATTCGCGCAAGAAAAGCAGATCCCTTCTGCG 728

Db 182 CYRGSTYKCMKAYTKRKRWTMTYYYYKSYMSNKKTWBMKATAYTKRMKTRKMTW 123
Qy 729 TGCTGGCAGTTTGAAGCTTCAACATGCTCTTGATTAATCCAGGAGCCATTTG 788
Db 122 CTGCMKCTTMMAGTMMYTRRYAKAKSKACTSTTTCMKYAKKMSYMSMS 63
Qy 789 CGCGCTGCTGATCTTCTTGATGCTGTTGACCATCCGCTACTTCTTTGGAAAAGCT 847
Db 62 MKMKGSMMWMTYTTTTYMMKSKMTYMSMSYAKKCMWTAKTYATYMTCTWGRKT 4

RESULT 11

US-09-621-976-18033
Sequence 18033, Application US/09621976
Patent No. 6639063
GENERAL INFORMATION:
APPLICANT: Dumas Milne Edwards, J.B.
APPLICANT: Jobert, S.
APPLICANT: Giordano, J.Y.
TITLE OF INVENTION: ESTs and Encoded Human Proteins.
FILE REFERENCE: GENSET.054PR2
CURRENT APPLICATION NUMBER: US/09/621,976
CURRENT FILING DATE: 2000-07-21
NUMBER OF SEQ ID NOS: 19335
SOFTWARE: Patent.pm
SEQ ID NO 18033
LENGTH: 474
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: 16
OTHER INFORMATION: n=a, g, c or t
US-09-621-976-18033

Query Match
Best Local Similarity 2.9%; Score 37; DB 4; Length 474;
Matches 49; Conservative 135; Mismatches 156; Indels 0; Gaps 0;

Qy 522 CGGCGAGCCCGCAGGCTGCTGCGCGCTGCGAGCTATCTCAATGAAATAGGCTTCACT 581
Db 13 YSGNCCYCCMAKRRSYGGRMYKSSCYKSCMCMKSTKSGSYKTTTWTW 72
Qy 582 CCTACTGGATTTGGCGGCTCTCACCGAGTGGGATGCGAGTGAATTCCTTTGAAA 641
Db 73 WTTTTCMAKRRMSGSKTTTMCSSKKTSCMAGRWKGKYYSRWYCYKAGCYM 132
Qy 642 TTAAGGCTCGAGTGGCTTGTCTCTCTCTTTGTCAGCTGACTTTGGATTCCTGCC 701
Db 133 WKRWYCS8CMYTKGGSMTTMMRRRKYKRYTKGKKKTTWMAAACYTWS 192
Qy 702 GAAGGAAAAGCAATCCCTCTCTGCTGCTGCGAGTGGAGCTTCAACATGCTCTTG 761
Db 193 YMMMRRAAAKTYTCMSKTMCAACCCMCMRRASCCMRSCMRBYTMMCYYY 252
Qy 762 TGTAAATTCAGGTCAGGCTTATTTGGCGGCTGCTGATTTTGGGCTGTGACCA 821
Db 253 MMYKGRMYMMWGGMMYEMTYMKKSMWKS CMWKAAMWARKTYTYTAMVYTYKR 312
Qy 822 TCCGCTACTCTCTCTGCGAAAAGCTGCTAATGACAAT 861
Db 313 MCCYMRKTTYCMWMSRWSMTARGAAMWMCYMY 352

RESULT 12

US-09-252-991A-3205/c
Sequence 3205, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136

Tue Apr 20 06:47:23 2004

us-10-608-504-6.inl

Page 5

IMMEDIATE SOURCE:
CLONE: PTZPc-Fls
US-08-232-463-14

Query Match 4.1%; Score 52.4; DB 1; Length 7218;
Best Local Similarity 3.6%; Pred. No. 2e-06;
Matches 14; Conservative 221; Mismatches 157; Indels 0; Gaps 0;

560 TCATGCAAAATAGCGTTCACTCTACGATTCGCGGCTCAGCCGAGTGGCGATC 619
1030 TCAGAGGAGCTGCTGATTTTATTAAGGCGCTGAGTGGCCCTTGTCTCTTCTC 679
620 GAGAGTGTATTCCTTTGAAATTAAGGCGCTGAGTGGCCCTTGTCTCTTCTC 679
1110 YY 1169
680 AGCTGACTTGGATTCCTCCGAGAAAGCAAGATCCCTCTCTGCTGCTGAGT 739
1170 YY 1229
740 TTGAGCTTCACTGCTCTCTGTAATTCAGGTCAGGCGCTTATTTGGCGGCTG 799
1230 YY 1289
800 ATCTTCTGGGCTCTGACATCCGATCTCTCTTCTTGGGAAAGGCTTAATGACA 859
1290 YY 1349
860 CTGATTTCTCTGATTCCTCTGTCGAGTATGTCAGATGTCATTACTTTGGCTCC 919
1350 YY 1409
920 GGGCGGTCGCTTCTTAATCTTAAGCCCTTA 951
1410 YYYYYYYYYYYYYYYYYYYYYYGTACCA 1441

RESULT 7

US-09-107-532A-1266
Sequence 1266, Application US/09107532A
Patent No. 6583275

GENERAL INFORMATION:

APPLICANT: Lynn A Doucette-Stamm and David Bush
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS

NUMBER OF SEQUENCES: 7310

CORRESPONDENCE ADDRESS:
ADDRESSEE: GENOME THERAPEUTICS CORPORATION
STREET: 100 Beaver Street
CITY: Waltham
STATE: Massachusetts
COUNTRY: USA
ZIP: 02354

COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660
COMPUTER: PC
OPERATING SYSTEM: <Unknown>
SOFTWARE: ASCII

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/107,532A
FILING DATE: 30-Jun-1998

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/085,598
FILING DATE: 14 May 1998

ATTORNEY/AGENT INFORMATION:
NAME: Ariniello, Pamela Denek
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-012

TELECOMMUNICATION INFORMATION:
TELEPHONE: (781)893-5007
TELEFAX: (781)893-8277

INFORMATION FOR SEQ ID NO: 1266:

SEQUENCE CHARACTERISTICS:
LENGTH: 768 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Enterococcus faecium
FEATURE:
NAME/KEY: misc feature
LOCATION: (B) LOCATION 1...768
SEQUENCE DESCRIPTION: SEQ ID NO: 1266:

US-09-107-532A-1266

Query Match 3.1%; Score 39.8; DB 4; Length 768;
Best Local Similarity 49.3%; Pred. No. 0.0072;
Matches 104; Conservative 0; Mismatches 107; Indels 0; Gaps 0;

207 TAAAACTCCCTCTGCTGACGTTTGGCATGTACCCGATGTTGCTTGTCT 266
92 TCAAGACAGTTACCTACCGTTTCGTTATTCGTTATGACTTGACATTGGATCG 151
267 TGGTATTCAATAGGCTAGCAATGGTGGCAGCCCACTGTTTCCGCGCTGATTTCG 326
152 TTGGAAAGCTGCCGATTCATTCACACTAGTCTGACGTTGATGCTGCTGCTATG 211
327 CGGCTCCACCGAATGCTGATACGCTCTGTTGGCGAGCGCCCTGGGCGCCA 386
212 CTGGTTCTGCCCATTTATACAGTCAGATGCTTGTACTACAGCCCATGCTTCCA 271
387 TCGGCTCACCATGCTGCTGTAACCTTCG 417
272 TGGTTCTGACCTTCTGATCAATTCGG 302

RESULT 8

US-09-252-991A-11376
Sequence 11376, Application US/09252991A
Patent No. 6551795

GENERAL INFORMATION:

APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A

CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788

PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190

PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142

SEQ ID NO 11376

LENGTH: 1038

TYPE: DNA

ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-11376

Query Match 3.1%; Score 39.2; DB 4; Length 1038;
Best Local Similarity 55.0%; Pred. No. 0.014;
Matches 77; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

309 TTCGCGCTGATTTTCGCGGCTTCACCGAAATGCTGTATGCTCTGTTGGGCG 368
473 TGTCCGATCGATCTCTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 532
369 CAGCGCCCTCGGCGCGCATCGCGCTCACCATGCTGTAATCTCGCCACGATTTCT 428
533 GTCGCGGCTTCTTTGATCATCTTTCACACACCTCTCTCTCTCTCTCTCTCT 592
429 ATGCGTTTCAATCCGCTG 448

Parent No. 6613545
 GENERAL INFORMATION:
 APPLICANT: KENNERKRECHT, NICOLE
 APPLICANT: SAHM, HERMANN
 APPLICANT: EGGELE, LOTHAR
 APPLICANT: PFEFFERLE, WALTER
 TITLE OF INVENTION: NUCLEOTIDE SEQUENCE CODING FOR THE EXPORT OF
 TITLE OF INVENTION: BRANCHED CHAIN AMINO ACIDS, PROCESS FOR THE
 TITLE OF INVENTION: ISOLATION THEREOF AND USE THEREOF
 FILE REFERENCE: 21123/265496/MAS
 CURRENT APPLICATION NUMBER: US/09/471.803A
 PRIOR FILING DATE: 1999-12-23
 PRIOR APPLICATION NUMBER: DE 199 51 708.8
 PRIOR FILING DATE: 1999-10-27
 NUMBER OF SEQ ID NOS: 12
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO: 4
 LENGTH: 324
 TYPE: DNA
 ORGANISM: Corynebacterium glutamicum
 FEATURE:
 NAME/KEY: CDS
 LOCATION: (1)..(324)
 OTHER INFORMATION: brne
 OTHER INFORMATION: ATCC14752
 US-09-471-803A-4

Query Match 25.5%; Score 324; DB 4; Length 324;
 Best Local Similarity 100.0%; Pred. No. 8.9e-99;
 Matches 324; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 853 ATGACACTGATTTCTCTCTGTAATTCCTCTGTTGCGACATATGCGAGTACTTACTTT 912
 DB 1 ATGACAACTGATTTCTCTCTGTAATTCCTCTGTTGCGAGATATGCGAGTACTTACTTT 60
 QY 913 GGGCTCCGGGCGGTTCCGTTTAAATCCTTAAGCCCTTACGATCAATTTGTGGGC 972
 DB 61 GGGCTCCGGGCGGTTCCGTTTAAATCCTTAAGCCCTTACGATCAATTTGTGGGC 120
 QY 973 AAAATGCGATGTGATGCGACAGGAAATCCTTGCCATTTTACCGCATCAACGTTTGGC 1032
 DB 121 AAAATGCGATGTGATGCGACAGGAAATCCTTGCCATTTTACCGCATCAACGTTTGGC 180
 QY 1033 AGCAATGCGATGTGATGCGACAGGAAATCCTTGCCATTTTACCGCATCAACGTTTGGC 1092
 DB 181 AGCAATGCGATGTGATGCGACAGGAAATCCTTGCCATTTTACCGCATCAACGTTTGGC 240
 QY 1093 GTGGCGCATCTTCTTGGCGGTGCGACGACCTTGTGAGCGTTGGCGCTGGCACTCTTT 1152
 DB 241 GTGGCGCATCTTCTTGGCGGTGCGACGACCTTGTGAGCGTTGGCGCTGGCACTCTTT 300
 QY 1153 TTTGTTGACTGCTGTAATCTTTTC 1176
 DB 301 TTTGTTGACTGCTGTAATCTTTTC 324

RESULT 5
 US-09-489-039A-3016
 Sequence 3016; Application US/09489039A
 Patent No. 6610836
 GENERAL INFORMATION:
 APPLICANT: Gary Breton et. al
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
 TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
 FILE REFERENCE: 2709.2004001
 CURRENT APPLICATION NUMBER: US/09/489.039A
 CURRENT FILING DATE: 2000-01-27
 PRIOR APPLICATION NUMBER: US 60/117,747
 PRIOR FILING DATE: 1999-01-29
 NUMBER OF SEQ ID NOS: 14342
 SEQ ID NO 3016
 LENGTH: 756
 TYPE: DNA

ORGANISM: Klebsiella pneumoniae
 US-09-489-039A-3016

Query Match 4.6%; Score 58.6; DB 4; Length 756;
 Best Local Similarity 52.2%; Pred. No. 3.4e-09;
 Matches 130; Conservative 0; Mismatches 119; Indels 0; Gaps 0;

QY 195 TCGCGAAGGTCTAAACCTTCCTGCTGACGATTTGGGCAATGACCGATGTATG 254
 DB 62 TCGCGAAGGTCTAAACCTTCCTGCTGACGATTTGGGCAATGACCGATGTATG 121
 QY 255 CGTTGCTCTTGTATTAATACGGTACGAATGGGAGAGCCCATGTTTCG 314
 DB 122 CTTGCGACTCAATGCTACCGGCTGGGCTTTACCCCGTGAAAGCTGTTCTTCT 181
 QY 315 GCGTATTTTCGGGCGCTCCACCGAATGCTGATCGCCCTGTTGTGGCGGACGCG 374
 DB 182 GCATTATTAAGCGCGCGCGACAGTGTGTATCACCGCATGTGCGCGCGACAGT 241
 QY 375 CCTGGCGCGCATCGGCTTACCAATGCTGTGTAATTCGCGACCATTTAGGT 434
 DB 242 CGGTGTGGTGGCGCGCTGACCGGTGATGCGATGATGTCGCGACGCTGTATGCGC 301
 QY 435 TTTCATTCC 443
 DB 302 CTTCACTGC 310

RESULT 6
 US-08-232-463-14
 Sequence 14; Application US/08232463
 Patent No. 5670367
 GENERAL INFORMATION:
 APPLICANT: DORNER, F.
 APPLICANT: SCHERFLINGER, F.
 APPLICANT: FALKNER, F. G.
 TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
 NUMBER OF SEQUENCES: 52
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Foley & Lardner
 STREET: 1800 Diagonal Road, Suite 500
 CITY: Alexandria
 STATE: VA
 COUNTRY: USA
 ZIP: 22313-0299
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/232.463
 FILING DATE:
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/07/935.313
 FILING DATE:
 APPLICATION NUMBER: EP 91 114 300.6
 FILING DATE: 26-AUG-1991
 ATTORNEY/AGENT INFORMATION:
 NAME: BENT, Stephen A.
 REGISTRATION NUMBER: 29,768
 REFERENCE/DOCKET NUMBER: 30472/114 IMMU
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (703)836-9300
 TELEFAX: (703)683-4109
 TELEX: 899149
 INFORMATION FOR SEQ ID NO: 14:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 7218 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear

```

Db 481 TTGGTTTGGCGCTTATCGACGAAGCCTACGACGTACGCGCGACGCGCGAGGCTG 540
Qy 541 GTGGCGGTGGGACTTATCTCAATGCAATAGCGTTTCACTCCCTACTGGGTATTCGGCGG 600
Db 541 GTGGCGGTGGGACTTATCTCAATGCAATAGCGTTTCACTCCCTACTGGGTATTCGGCGG 600
Qy 601 TCTACCGGAGTGGCGATCGCAGAGTTGATCTCTTTTGAATTAAGGAGCTCGAGTTCG 660
Db 601 TCTACCGGAGTGGCGATCGCAGAGTTGATCTCTTTTGAATTAAGGAGCTCGAGTTCG 660
Qy 661 CTTTGTCTCTCTTTTGTACGCTGATCTTGAATTCCTCCGAAAGCAAAAGCATGCC 720
Db 661 CTTTGTCTCTCTTTTGTACGCTGATCTTGAATTCCTCCGAAAGCAAAAGCATGCC 720
Qy 721 TTCTGTGTGTGTGGAGTTGAGGTTGACCATGCTGTTGTGTAATTCAGGTACAGG 780
Db 721 TTCTGTGTGTGTGGAGTTGAGGTTGACCATGCTGTTGTGTAATTCAGGTACAGG 780
Qy 781 CCTATTGCGCGCGCTGATCTTCTTGGGTCTGTGAACATCGGTAATCTTCTTGGG 840
Db 781 CCTATTGCGCGCGCTGATCTTCTTGGGTCTGTGAACATCGGTAATCTTCTTGGG 840
Qy 841 AAAGGCTCTAAATGACAACTGATTTCTCCCTGATCTCCCTGTTGTGCGAGTATGCA 900
Db 841 AAAGGCTCTAAATGACAACTGATTTCTCCCTGATCTCCCTGTTGTGCGAGTATGCA 900
Qy 901 GTCAATTACTTTTGGCGCTCGGCGGTTCCGTTCTTAATCCTTAAGCCCTACGTAA 960
Db 901 GTCAATTACTTTTGGCGCTCGGCGGTTCCGTTCTTAATCCTTAAGCCCTACGTAA 960
Qy 961 CAATTGTTGGGCAAAATGGCGATGTGATGCGAGAGAAATCTTGCAATTGACGCA 1020
Db 961 CAATTGTTGGGCAAAATGGCGATGTGATGCGAGAGAAATCTTGCAATTGACGCA 1020
Qy 1021 TCAAGCTTTGCGAGCAATGCGATAGATCTGAAGACTTAACCTTGGTCTATGGCGGT 1080
Db 1021 TCAAGCTTTGCGAGCAATGCGATAGATCTGAAGACTTAACCTTGGTCTATGGCGGT 1080
Qy 1081 GCGATTACAGTGTGGCGCATCTTTTGGCGGTGACGACACTTGTGAGCGGTGCGCT 1140
Db 1081 GCGATTACAGTGTGGCGCATCTTTTGGCGGTGACGACACTTGTGAGCGGTGCGCT 1140
Qy 1141 GCGACCATCGTTTGTGTGACTGCGTGAATCTTTTCTAAATCGATTAATTAACAAAT 1200
Db 1141 GCGACCATCGTTTGTGTGACTGCGTGAATCTTTTCTAAATCGATTAATTAACAAAT 1200
Qy 1201 CGGCAATGCGCTCAATTTGAAGGAGTGGGATTTTGAAGAACCTAGAAAAGGCTTAA 1260
Db 1201 CGGCAATGCGCTCAATTTGAAGGAGTGGGATTTTGAAGAACCTAGAAAAGGCTTAA 1260
Qy 1261 CAGACAGGCGCT 1271
Db 1261 CAGACAGGCGCT 1271

```

```

RESULT 3
US-09-471-803A-2
; Sequence 2, Application US/09471803A
; Patent No. 6613545
; GENERAL INFORMATION:
; APPLICANT: KENNERKNECHT, NICOLE
; APPLICANT: SAHM, HERMAN
; APPLICANT: EGGELING, LOTMAR
; APPLICANT: PEEFERLE, WALTER
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCE CODING FOR THE EXPORT OF
; TITLE OF INVENTION: BRANCHED CHAIN AMINO ACIDS. PROCESS FOR THE
; TITLE OF INVENTION: ISOLATION THEREOF AND USE THEREOF
; FILE REFERENCE: 21123/263496/MAS
; CURRENT APPLICATION NUMBER: US/09/471,803A
; CURRENT FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: DE 199 51 708-8
; PRIOR FILING DATE: 1999-10-27

```

```

; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 2
; LENGTH: 753
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(753)
; OTHER INFORMATION: brnF
; OTHER INFORMATION: ATCC14752
; US-09-471-803A-2

Query Match      59.2%; Score 753; DB 4; Length 753;
Best Local Similarity 100.0%; Pred.No. 6,66-243;
Matches 753; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 101 GTGCAAAAACGCAAGATTCATTCAAGCTTGAGGTGTGCGCATCCAAAGCAGCGCTG 160
Db 1 GTGCAAAAACGCAAGATTCATTCAAGCTTGAGGTGTGCGCATCCAAAGCAGCGCTG 60
Qy 161 GAACGAGATGATTAAGTTATCGGCTACGAATCGCGCAAGTCTAATAAAGCTCCCT 220
Db 61 GAACGAGATGATTAAGTTATCGGCTACGAATCGCGCAAGTCTAATAAAGCTCCCT 120
Qy 221 GCTGCAAGTTTGGGCAATGACCCGATGATGATGATGATGATGATGATGATGATGAT 280
Db 121 GCTGCAAGTTTGGGCAATGACCCGATGATGATGATGATGATGATGATGATGATGAT 180
Qy 281 GGCTAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 340
Db 181 GGCTAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 240
Qy 341 ATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 400
Db 241 ATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300
Qy 401 TTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 460
Db 301 TTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 360
Qy 461 AACCCATGCGCGGTTCTATTGCGTTTGGCGCTTATGACGAAGCTGACGACT 520
Db 361 AACCCATGCGCGGTTCTATTGCGTTTGGCGCTTATGACGAAGCTGACGACT 420
Qy 521 GCGGCGAGCGCGGAGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCG 580
Db 421 GCGGCGAGCGCGGAGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCG 480
Qy 581 TCTACTGAGTATTCGCGCGGTCTACCGGAGTGGCGATGCGAGATGATTCCTTTGAA 640
Db 481 TCTACTGAGTATTCGCGCGGTCTACCGGAGTGGCGATGCGAGATGATTCCTTTGAA 540
Qy 641 ATTAAGGCGCTGAGTTCGCGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 700
Db 541 ATTAAGGCGCTGAGTTCGCGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 600
Qy 701 CGAAGCAAAAAGCAGATCCCTTCTCTGCTGCTGCGAGTTTGAAGCTTCAACATGCTCT 760
Db 601 CGAAGCAAAAAGCAGATCCCTTCTCTGCTGCTGCGAGTTTGAAGCTTCAACATGCTCT 660
Qy 761 GTGTATTCAGAGTCAAGGCGCTTATTTGGCGGCTGCTGATCTTCTTGGGTCTGTGACC 820
Db 661 GTGTATTCAGAGTCAAGGCGCTTATTTGGCGGCTGCTGATCTTCTTGGGTCTGTGACC 720
Qy 821 ATCCGATCTTCTTGGGAAAGGCTGCTAA 853
Db 721 ATCCGATCTTCTTGGGAAAGGCTGCTAA 753

```

```

RESULT 4
US-09-471-803A-4
; Sequence 4, Application US/09471803A

```

```

Db      181  TCGGGGCTACGAAATCGCGCAGGCTCTAAAAACCTCCCTTGTGAGGTTGGGAGTGA 240
QY      241  CCCGATGGTATGCGTTGGTGTCTCTGGTATTTCAATACGGCTACGATGGTGGGACG 300
Db      241  CCCGATGGTATGCGTTGGTGTCTCTGGTATTTCAATACGGCTACGATGGTGGGACG 300
QY      301  CCCACTGTTTTCCGGCTGATTTTGGCGGGCTCCACCGAAATGCTGTGATCGCCCTGT 360
Db      301  CCCACTGTTTTCCGGCTGATTTTGGCGGGCTCCACCGAAATGCTGTGATCGCCCTGT 360
QY      361  TGTGGGCGGAGGCGCCCTGGGGCGCATGGGCTCACCATATGCTGTGAATTCGCGCA 420
Db      361  TGTGGGCGGAGGCGCCCTGGGGCGCATGGGCTCACCATATGCTGTGAATTCGCGCA 420
QY      421  CGTATTCATAGCGTTTCAATCCCGCTGAGTGTGTAATAAACCCCATTTGCCGTTCTA 480
Db      421  CGTATTCATAGCGTTTCAATCCCGCTGAGTGTGTAATAAACCCCATTTGCCGTTCTA 480
QY      481  TTGCGTTTCCGGCTTATGAGAGAACCTACGCACTGGGCGAGGCGCGAGGCTG 540
Db      481  TTGCGTTTCCGGCTTATGAGAGAACCTACGCACTGGGCGAGGCGCGAGGCTG 540
QY      541  GTGCGGCTGGCGCATTAATCTCAATGCAATAGCGTTCACTGCGTATTCGGCGG 600
Db      541  GTGCGGCTGGCGCATTAATCTCAATGCAATAGCGTTCACTGCGTATTCGGCGG 600
QY      601  TCTCACCGGAGTGGCGCATGGCAGACTGATTCCTTTGAATTAAGGCGCTGAGTGGC 660
Db      601  TCTCACCGGAGTGGCGCATGGCAGACTGATTCCTTTGAATTAAGGCGCTGAGTGGC 660
QY      661  CCTTGTCTCTCTTTGTGACGCTGACTTGTGAATTCCTGCGCAGAAAGACAGATCCC 720
Db      661  CCTTGTCTCTCTTTGTGACGCTGACTTGTGAATTCCTGCGCAGAAAGACAGATCCC 720
QY      721  TTCTCTGCTGCTGGCGGTTTGAAGCTTCACATTCCTTGTGTGATTCGAGTCAAGC 780
Db      721  TTCTCTGCTGCTGGCGGTTTGAAGCTTCACATTCCTTGTGTGATTCGAGTCAAGC 780
QY      781  CCTATTGGCGGCGCTGATCTTCTTGGGCTGTGTGACCATCCGCTACTTCTTCTTGGG 840
Db      781  CCTATTGGCGGCGCTGATCTTCTTGGGCTGTGTGACCATCCGCTACTTCTTCTTGGG 840
QY      841  AAAAGCTGCTAAATGACAACTGATTTCTCTGTATTCCTTGTGTGCAAGTATGCA 900
Db      841  AAAAGCTGCTAAATGACAACTGATTTCTCTGTATTCCTTGTGTGCAAGTATGCA 900
QY      901  GTGATTAATCTTTGGCGTCCGGGCGGTTCCGTTCTTAATCCTTAAGCCCTTACGTAATCA 960
Db      901  GTGATTAATCTTTGGCGTCCGGGCGGTTCCGTTCTTAATCCTTAAGCCCTTACGTAATCA 960
QY      961  CAATTTGTGGGCAAAATGCGAGTGTGGATCCAGCAGAAATCCTTCCATTTTGAACGCA 1020
Db      961  CAATTTGTGGGCAAAATGCGAGTGTGGATCCAGCAGAAATCCTTCCATTTTGAACGCA 1020
QY      1021  TCAACGTTTGGCGCAATGCGATAGATCTGAAGACTTAACTTGTGTGATTCGCGCT 1080
Db      1021  TCAACGTTTGGCGCAATGCGATAGATCTGAAGACTTAACTTGTGTGATTCGCGCT 1080
QY      1081  GCGATTAACAGTGTGGCGCATCTTCTTGGCGGTGACGACCTTGTGAGCGTGGCGCT 1140
Db      1081  GCGATTAACAGTGTGGCGCATCTTCTTGGCGGTGACGACCTTGTGAGCGTGGCGCT 1140
QY      1141  GCGACCAATGCTTTTGTGAGACTGTGAATCTTTCTTAAACTGCAATTAACAATAAT 1200
Db      1141  GCGACCAATGCTTTTGTGAGACTGTGAATCTTTCTTAAACTGCAATTAACAATAAT 1200
QY      1201  CCGCATGCGCTCAATTTGAAGGGAGTGGGATTTTGAAGAACTTGAAGAAAGGCTTAAG 1260
Db      1201  CCGCATGCGCTCAATTTGAAGGGAGTGGGATTTTGAAGAACTTGAAGAAAGGCTTAAG 1260
QY      1261  CAGACAGCGCT 1271

```

```

Db      1261  CAGACAGCGCT 1271

RESULT 2
US-09-471-803A-6
; Sequence 6, Application US/09471803A
; Patent No. 6613545
; GENERAL INFORMATION:
; APPLICANT: KENNEDY, NICOLE
; APPLICANT: SAHM, HERMAN
; APPLICANT: EGBELING, LOTAR
; APPLICANT: PEEFERLE, WALTER
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCE CODING FOR THE EXPORT OF
; TITLE OF INVENTION: BRANCHED CHAIN AMINO ACIDS, PROCESS FOR THE
; TITLE OF INVENTION: ISOLATION THEREOF AND USE THEREOF
; FILE REFERENCE: 21123/265496/MAS
; CURRENT APPLICATION NUMBER: US/09/471, 803A
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: DE 199 51 708.8
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Patent Ver. 2.1
; SEQ ID NO 6
; LENGTH: 1271
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
; FEATURE:
; NAME/KEY: gene
; LOCATION: (101)..(853)
; OTHER INFORMATION: DnF
; NAME/KEY: gene
; LOCATION: (853)..(1176)
; OTHER INFORMATION: DnF
; OTHER INFORMATION: ATCC13032
US-09-471-803A-6

Query Match      100.0%; Score 1271; DB 4; Length 1271;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1271; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1  GCGCATCAATGCAATGATCTTCAATATTTGACATATAGCTAGTGTGAGTGGCGAAAC 60
Db      1  GCGCATCAATGCAATGATCTTCAATATTTGACATATAGCTAGTGTGAGTGGCGAAAC 60
QY      61  TGGCAACAAATCTACCGGCAATTTGTGATGATTTGTAGTGTGCAAAAAACGCAAGAT 120
Db      61  TGGCAACAAATCTACCGGCAATTTGTGATGATTTGTAGTGTGCAAAAAACGCAAGAT 120
QY      121  TCATTCAAGCTGAGAGTGTGCGCATCAAGGAGCCCTGGAACCAAGATATAAGGTTA 180
Db      121  TCATTCAAGCTGAGAGTGTGCGCATCAAGGAGCCCTGGAACCAAGATATAAGGTTA 180
QY      181  TGGGCGCTACGAAATTCGCGCAAGGTCTPAAAACTCCTCTTGTGCGAGTTTGGCATGTA 240
Db      181  TGGGCGCTACGAAATTCGCGCAAGGTCTPAAAACTCCTCTTGTGCGAGTTTGGCATGTA 240
QY      241  CCCGATGGTATGCGTTGGTGTCTCTGGTATTTCAATACGGCTACGATGGTGGGACG 300
Db      241  CCCGATGGTATGCGTTGGTGTCTCTGGTATTTCAATACGGCTACGATGGTGGGACG 300
QY      301  CCCACTGTTTTCCGGCTGATTTTGGCGGGCTCCACCGAAATGCTGTGATCGCCCTGT 360
Db      301  CCCACTGTTTTCCGGCTGATTTTGGCGGGCTCCACCGAAATGCTGTGATCGCCCTGT 360
QY      361  TGTGGGCGGAGGCGCCCTGGGGCGCATGGGCTCACCATATGCTGTGAATTCGCGCA 420
Db      361  TGTGGGCGGAGGCGCCCTGGGGCGCATGGGCTCACCATATGCTGTGAATTCGCGCA 420
QY      421  CGTATTCATAGCGTTTCAATCCCGCTGAGTGTGTAATAAACCCCATTTGCCGTTCTA 480
Db      421  CGTATTCATAGCGTTTCAATCCCGCTGAGTGTGTAATAAACCCCATTTGCCGTTCTA 480
QY      481  TTGCGTTTCCGGCTTATGAGAGAACCTACGCACTGGGCGAGGCGCGAGGCTG 540

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd

OM nucleic - nucleic search, using sw model

```
Run on:      April 19, 2004, 10:57:33 ; Search time 107.018 Seconds
              (without alignments)
              6590.898 Million cell updates/sec
```

```

Title:          US-10-608-504-6
Perfect score:  1271
Sequence:       1 gcgcgcatcaatggaatctag.....aggcttaagcagacagcgct 1271

```

Scoring table: IDENTITY -noc
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

```
Minimum DB seq length: 0
Maximum DB seq length: 20000000000
```

```
Post-processing: Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45 summaries
```

```
Database :
Issued_Patents NA:*
1: /cgnr2_6/prodata/2/ina/5A.COMB.seq:*
2: /cgnr2_6/prodata/2/ina/5B.COMB.seq:*
3: /cgnr2_6/prodata/2/ina/6A.COMB.seq:*
4: /cgnr2_6/prodata/2/ina/6B.COMB.seq:*
5: /cgnr2_6/prodata/2/ina/PCUTUS.COMB.seq:*
6: /cgnr2_6/prodata/2/ina/bactflies.seq:
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1271	100.0	1271	4	US-09-471-803A-1	Sequence 1, Appl1
2	1271	100.0	1271	4	US-09-471-803A-6	Sequence 6, Appl1
3	753	59.2	753	4	US-09-471-803A-2	Sequence 2, Appl1
4	324	25.5	324	4	US-09-471-803A-4	Sequence 4, Appl1
5	58.6	4.6	756	4	US-09-489-039A-3016	Sequence 3016, Ap
6	52.4	4.1	7218	1	US-08-332-463-14	Sequence 14, Appl
7	39.8	3.1	768	4	US-09-107-532A-1266	Sequence 1266, Ap
8	39.2	3.1	1038	4	US-09-252-991A-11376	Sequence 11376, A
9	39.2	3.1	1086	4	US-09-252-991A-11418	Sequence 11418, A
10	39	3.1	399	4	US-09-621-976-8976	Sequence 8976, Ap
11	37	2.9	474	4	US-09-621-976-18033	Sequence 18033, Ap
12	36.8	2.9	1140	4	US-09-252-991A-3205	Sequence 3205, Ap
13	36.8	2.9	2061	4	US-09-252-991A-2761	Sequence 2761, Ap
14	36.4	2.9	751	3	US-09-020-956-12	Sequence 12, Appl
15	36.4	2.9	751	3	US-09-020-607-12	Sequence 12, Appl
16	36.4	2.9	751	4	US-09-439-313-12	Sequence 12, Appl
17	36.4	2.9	751	4	US-09-352-616A-12	Sequence 12, Appl
18	36.4	2.9	751	4	US-09-232-149A-12	Sequence 12, Appl
19	36.4	2.9	751	4	US-09-159-812-12	Sequence 12, Appl
20	36.4	2.9	751	4	US-09-636-215-12	Sequence 12, Appl
21	36.4	2.9	751	4	US-09-665-166A-12	Sequence 12, Appl
22	36.4	2.9	751	4	US-09-115-453-12	Sequence 12, Appl
23	36.4	2.9	751	4	US-09-688-489-12	Sequence 12, Appl
24	35.2	2.8	430	4	US-09-621-976-16656	Sequence 16656, A
25	35.2	2.8	801	4	US-09-338-352-2399	Sequence 2399, Ap
26	35.2	2.8	4403765	3	US-09-103-840A-2	Sequence 2, Appl
27	35.2	2.8	4411529	3	US-09-103-840A-1	Sequence 1, Appl

28	34.8	2.7	867	4	US-09-489-039A-5047	Sequence 5047, Appl1
29	34.6	2.7	866	2	US-08-875-034A-1	Sequence 1, Appl1
30	34.4	2.7	744	4	US-09-489-039A-1307	Sequence 1307, Appl1
31	34.4	2.7	1234	4	US-09-426-965-12	Sequence 22, Appl1
32	34.4	2.7	1229	4	US-09-426-965-74	Sequence 74, Appl1
33	34.4	2.7	1831	4	US-09-226-965-06	Sequence 66, Appl1
34	34.2	2.7	234	4	US-09-252-991A-13729	Sequence 13729, A
35	34.2	2.7	678	4	US-09-252-991A-13832	Sequence 13832, A
36	33.4	2.6	1875	3	US-09-422-869-21	Sequence 21, Appl1
37	33.4	2.6	2949	3	US-08-433-522A-1	Sequence 1, Appl1
38	33.4	2.6	2949	3	US-09-135-166-1	Sequence 1, Appl1
39	33.4	2.6	2949	3	US-08-942-046-1	Sequence 1, Appl1
40	33.4	2.6	2950	3	US-08-933-522A-5	Sequence 5, Appl1
41	33.4	2.6	2950	3	US-09-135-166-5	Sequence 5, Appl1
42	33.4	2.6	2950	3	US-08-942-046-5	Sequence 5, Appl1
43	33.4	2.6	2984	3	US-08-433-522A-3	Sequence 3, Appl1
44	33.4	2.6	2984	3	US-09-135-166-3	Sequence 3, Appl1
45	33.4	2.6	2984	3	US-08-942-046-3	Sequence 3, Appl1

ALIGNMENTS

```

RESULT 1
US-09-471-803A-1
; Sequence 1, Application US/09471803A
; Patent No. 6613545
; GENERAL INFORMATION:
; APPLICANT: KENNERKNECHT, NICOLE
; APPLICANT: SAHM, HERMANN
; APPLICANT: EGGSEILING, LOTAR
; APPLICANT: PFEFFERLE, WALTER
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCE CODING FOR THE EXPORT OF
; TITLE OF INVENTION: BRANCHED CHAIN AMINO ACIDS, PROCESS FOR THE
; TITLE OF INVENTION: ISOLATION THEREOF AND USE THEREOF
; FILE REFERENCE: 21123/265496/MAS
; CURRENT FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: DE 199 51 708.8
; PRIOR FILING DATE: 1999-10-27
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 1
; LENGTH: 1271
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
; FEATURES:
; NAME/KEY: gene
; LOCATION: (101)..(853)
; OTHER INFORMATION: brnF
; NAME/KEY: gene
; LOCATION: (853)..(1176)
; OTHER INFORMATION: brnE
; OTHER INFORMATION: ATCC14752
; US-09-471-803A-1

Query Match      100.0%; Score 1271; DB 4; Length 1271;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1271; Conservative 0; Mismatches 0; Indels 0; Gaps 0

QY      1 GGGCGATCAATGGAATCTAGACTTCAATATTTGCAACAATAAGCTAATTGAGTGGCCGAAC 60
DB      1 GGGCGATCAATGGAATCTAGACTTCAATATTTGCAACAATAAGCTAATTGAGTGGCCGAAC 60
QY      61 TGGCAACAAAACCTACCCGGCAATTGTGTGATGATTGTGTGCAAAAAAACCAGAAGAT 120
DB      61 TGGCAACAAAACCTACCCGGCAATTGTGTGATGATTGTGTGCAAAAAAACCAGAAGAT 120
QY      121 TCATTCAAGCTGTGAGTGTGTGCCATCCAAAGGACCCCTTGCAACCAAGATTAAGTTTA 180
DB      121 TCATTCAAGCTGTGAGTGTGTGCCATCCAAAGGACCCCTTGCAACCAAGATTAAGTTTA 180
QY      181 TCGGGCGCTACGAATACTGGGCGCAAGTCTAAAAACCTCCCTGTGTGAGGTTTGGGCAATGTA 240

```

```
Db 233 TCATGGTAATTTCCGACATTTTCTACGCTCACCTTCCACGCCACCGCATCAAGT 232
QY 462 ACCCATTTCCCGTTTCTATTGGTTTGGCTTATGACGAAGCTTAGCACTCACTG 521
Db 293 CCGGCGCGCGCCGCGCTATTCCACCTACGCGCTTACGACGAGCTCTACGCCATCGTGT 352
QY 522 CGGCGAGGCCCGCGAGCTGTGCGCGTG---GCCACTTATCTCAATGGAATAGCTTTC 578
Db 353 CAGCCGCCCACTGCGCATATCACTGACACGCGGTCCTTACCGTTCAATTTTGTGCC 412
QY 579 ACTCTACTGGGATTTGCGCGCTCTCACCGAGATGCGCATGCAAGTTGATTCCTTTG 638
Db 413 AAGCTCTGTGGGTATCCAGGAATTTATGGCGCTTGTGTGTCGAAGTGTGCCGATG 472
QY 639 AAATTAAAGGCTCGAGTTGCGCCCTTGTCTCTCTTTGTCAAGCTGACTTTGATTCCT 698
Db 473 ATCTAAAGGATGATTTTGCCCTGACCGCGCTGTGTGTGCTGTGCGGTGGAGGAT 532
QY 699 GCCGAACGAAAAGCAGATCCCTCTCTGCTGCTGCAAGTTTGAGCTTCAACATTTGCTC 758
Db 533 TCAAAATATACAGATATTATGCGCATATTGCGGTGATATGGCTCTGTTCCG 592
QY 759 TTGTGTAATTCAGAGTCAGGCCCTATTGCGGCGCTGTGATCTTCTTGGGTCTGTGA 818
Db 593 GTTTGTGCGCGCCGAGCAGATGCTGTATGCTTTGACCACGTACTTTTGAATCCTTC 652
QY 819 CCATCCGGTACTTCTTC 835
Db 653 TTCTCCGCGTCCGCTTC 669
```

Search completed: April 19, 2004, 11:42:05
Job time : 593.9 secs

XX 16-DEC-1999; 99JP-00377484.
 PR 07-APR-2000; 2000JP-00159162.
 PR 03-AUG-2000; 2000JP-00280988.
 XX
 PA (KYO) KYOMA HAKKO KOGYO KK.
 XX Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;
 PI Tateishi N, Senoh A, Ikeda M, Ozaki A;
 XX WPI: 2001-376931/40.
 DR P-PSDB: AAG93154.
 XX
 PT Novel polynucleotides derived from Corynebacterium bacteria, for identifying
 PT mutation point of a gene, measuring expression of a gene, analyzing
 PT expression profile or pattern of a gene and identifying homologous gene.
 XX
 PS Claim 8; SEQ ID NO 3408; 246pp + Sequence Listing; English.
 CC The present invention provides a number of nucleotide and protein
 CC sequences, from the Corynebacterium glutamicum glutamicum. These
 CC are useful for identifying the mutation point of a gene derived from a
 CC mutant of corynebacterium glutamicum, measuring expression amount and analyzing
 CC the expression profile or expression pattern of a gene derived from
 CC Corynebacterium glutamicum, and identifying a homologue of a gene derived from
 CC corynebacterium glutamicum. Corynebacterium bacteria are useful for producing amino
 CC acids, nucleic acids, vitamins, saccharides and organic acids,
 CC particularly L-lysine. The present sequence is a nucleic acid described
 CC in the exemplification of the invention. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from the European Patent Office
 XX
 SQ Sequence 711 BP; 123 A; 190 C; 198 G; 200 T; 0 U; 0 Other;
 Query Match 10.0%; Score 127.4; DB 5; Length 711;
 Best Local Similarity 51.5%; Pred. No. 9e-30;
 Matches 318; Conservative 0; Mismatches 296; Indels 3; Gaps 1;
 QY 222 CTGCAAGTTTGGGAGTACCGGATTTGGTATGCTCTGTTATTCATACG 281
 DB 53 CGGTGGTTTGGGAGTATTCGCGTGGTGGCTTTGGCTTTGATGATGATGATGATGAT 112
 QY 282 GCTACGAATGGTGGGAGCCGCACTGTTTCCGCGCTGATTTTGGCGGCTCCACCGAAA 341
 DB 113 GTTTCGCTGGTGGTGGAGCGCCGATTTTCTCTTGATGATGATGATGATGATGAT 172
 QY 342 TGGTGTGATACGCGCTGTTGGGCGAGCGCCGCTGGGCGCATGGCGCTCCACCAT 401
 DB 173 TTTCTGCAATCGGCAATGATACCGCAGATGATGATGATGATGATGATGATGATGAT 232
 QY 402 TGGTGTGAACTTCCGCGCAGATTTCTATGCGTTTTCATCCCGCTGCATGTGTCAAA 461
 DB 223 TGATGTGATTTCCGCGCAGATTTTCTAGGCTCACTTCCACGCGCAGCATCAAGT 292
 QY 462 ACCCATGCGCGCTTTCTATTTGGTTTGGCGCTTATGAGAGGAGGAGGAGGAGGAGG 521
 DB 293 CGGCGCGCGCGCGCGCTTATTCACCTTACGCGCTTACGAGAGTCTTACGCGCATGCTGT 352
 QY 522 CGGCGAGGCGCGCAGGCTGTGCGGCGT--GCGACTTATCTCATGCAATGCAATAGCGTTTC 578
 DB 353 CAGCGCGCGCAGCTGTGCGGATTCAGTGGAGCGCGGCTGATTCAGTCAAAATTTGTGCC 412
 QY 578 ACTCTACCTGGGATTTGGCGGCTTCAAGGAGTGGCGATGCGAGATGATTTCTTTTG 638
 DB 413 AAGCTCTGGGATTTTCCAGGAATTTATGGCGCTTGGTGTGCAAGTGTGCGCGCATG 472
 QY 639 AAATTAAGGCGCTTCAGTTGCGCTTGTCTCTCTTTGTCAAGCTTGAATTCCT 698
 DB 473 ACTTAAGGAGATGATTTTGGCGCGAGCGCGCTGTTTGTGCTGGGATGAGGAGCAT 532
 QY 699 GCGGAAGCAAAAGCAGATCCCTTCTCTGCTGCTGCGAGTTTGAAGCTTCAACATTTGCTC 758
 DB 533 TCAAAATTAACAGGATTTATGCTGCGCATTTTGGCGGATGATGCTGTGTTCCG 592

QY 759 TTGTGTAATTCAGATGAGCCCTTATTTGGCGCGCTGCTGATCTTTGGTCTGTGA 818
 DB 593 GTTTTGTGGCGCCGCGAGAGATGTGTATGCTTTGACCGAGTATTTGATCTTC 652
 QY 819 CCATCGGATCTTCTTC 835
 DB 653 TTCTCGCGCTCCGCTTC 669
 RESULT 15
 ID ACA01294
 XX ACA01294 standard; DNA; 711 BP.
 AC ACA01294;
 XX
 DT 03-JUN-2003 (first entry)
 DE C. glutamicum derived ORF SEQ ID 1285.
 XX
 KM Corynebacterium; nucleic acid array; fermentation; culture; ds.
 XX
 OS Corynebacterium glutamicum.
 XX DE10128510-A1.
 XX PD 19-DEC-2002.
 XX PF 13-JUN-2001; 2001DE-01028510.
 XX PR 13-JUN-2001; 2001DE-01028510.
 XX PA (DEGS) DEGUSA AG.
 PI Farwick M, Moeskel B, Pfefferle W, Bathe B, Hutmacher K;
 DR WPI: 2003-279970/28.
 XX
 PT New nucleic acid array useful for monitoring mRNA expression of
 PT Corynebacterium glutamicum during fermentation, comprising nucleic acid
 PT from Corynebacterium glutamicum.
 XX
 PS Claim 1; Page 456; 709pp; German.
 XX
 CC This invention describes a novel nucleic acid array involving
 CC Corynebacterium glutamicum polynucleotides. The arrays are used to
 CC analyse C. glutamicum, particularly for monitoring a fermentation process
 CC to determine expression levels of C. glutamicum cellular mRNA. Such
 CC monitoring particularly differentiates between expression levels of
 CC different strains of C. glutamicum and allows the adjustment of different
 CC culture and fermentation conditions. ACA0010-ACA02188 represent C.
 CC glutamicum derived polynucleotides described in the disclosure of the
 CC invention
 XX
 SQ Sequence 711 BP; 123 A; 190 C; 198 G; 200 T; 0 U; 0 Other;
 Query Match 10.0%; Score 127.4; DB 7; Length 711;
 Best Local Similarity 51.5%; Pred. No. 9e-30;
 Matches 318; Conservative 0; Mismatches 296; Indels 3; Gaps 1;
 QY 222 CTGCAAGTTTGGGAGTACCGGATTTGGTATGCTCTGTTATTCATACG 281
 DB 53 CGGTGGTTTGGGAGTATTCGCGTGGTGGCTTTGGCTTTGATGATGATGATGATGAT 112
 QY 282 GCTACGAATGGTGGGAGCCGCACTGTTTCCGCGCTGATTTTGGCGGCTCCACCGAAA 341
 DB 113 GTTTCGCTGGTGGTGGAGCGCCGATTTTCTCTTGATGATGATGATGATGATGAT 172
 QY 342 TGGTGTGATACGCGCTGTTGGGCGAGCGCCCTGGGCGCGCATGCGCTCACCAT 401
 DB 173 TTTCTGCAATCGGCAATGATACCGCAGATGATGATGATGATGATGATGATGATGAT 232
 QY 402 TGGTGTGAACTTCCGCGCAGATTTCTATGCGTTTTCATTTCCGCTGCATGTGTCAAA 461


```

FT FT      -10_signal      /*tag= a
FT FT      88..93
FT FT      /*tag= b
FT FT      151..615
FT FT      /*tag= c
FT FT      /product= "ltp"
XX XX
XX XX      EP1090993-A1.
XX XX
XX XX      11-APR-2001.
XX XX
XX XX      29-SEP-2000; 2000EP-00121159.
XX XX
XX XX      05-OCT-1999; 99DE-01047792.
XX XX
XX XX      (DEGS ) DEGUSA-HUELS AG.
XX XX
XX XX      Moeckel B, Pfeifferle W, Puehler A, Kalinowski J, Bathe B;
XX XX      WPI, 2001-292927/31.
XX XX      P-PSDB; AAB70881.
XX XX
XX XX      New ltp gene from coryneform bacteria, used to prepare transformants with
XX XX      increased synthesis of amino acids, particularly lysine and isoleucine.
XX XX
XX XX      Claim 4; Page 14-15; 22pp; German.
XX XX
XX XX      This invention describes a novel isolated nucleic acid (I) from
XX XX      coryneform bacteria used for the fermentative production of selected L-
XX XX      amino acids, by fermenting the amino acid-producing coryneform in which
XX XX      at least the ltp gene has been weakened or amplified, then isolating
XX XX      amino acids that have accumulated in the medium or cells. (I) is used to
XX XX      transform coryneforms for production of L-amino acids, specifically
XX XX      lysine and isoleucine, which are used in medicine and particularly as
XX XX      animal feed supplement. It may also be used as probes and primers for
XX XX      isolating related sequences. Regulating expression of (I) improves
XX XX      production of amino acids, especially of L-lysine. This sequence encodes
XX XX      the Corynebacterium glutamicum ltp protein which is used in the method
XX XX      described in the invention
XX XX
XX XX      Sequence 715 BP; 166 A; 173 C; 178 G; 198 T; 0 U; 0 Other;
XX XX
XX XX      Query Match      14.6%; Score 185; DB 4; Length 715;
XX XX      Best Local Similarity 100.0%; Pred. No. 2.9e-48;
XX XX      Matches 185; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX XX
XX XX      1 GCGGATCAATGGAATCTAGCTCATATATTGCAATAGCTAGTTGAGTCCGCAAC 60
XX XX      185 GCGGATCAATGGAATCTAGCTCATATATTGCAATAGCTAGTTGAGTCCGCAAC 126
XX XX
XX XX      61 TGGCAACAAACTACCCGGCAATTGTGTGATGATTTGATGTCGCAAAAAACGCAAGAT 120
XX XX      125 TGGCAACAAACTACCCGGCAATTGTGTGATGATTTGATGTCGCAAAAAACGCAAGAT 66
XX XX
XX XX      121 TCATTCAAGCTTGAGGTGTGCGATCCAGGAGCCCTGGAACGAGATTAAGGTTA 180
XX XX      65 TCATTCAAGCTTGAGGTGTGCGATCCAGGAGCCCTGGAACGAGATTAAGGTTA 6
XX XX
XX XX      181 TCGGC 185
XX XX      5 TCGGC 1
XX XX
XX XX      RESULT 13
XX XX      ID ACA01968 standard; DNA; 177 BP.
XX XX      ACA01968;
XX XX      AC
XX XX      04-JUN-2003 (first entry)
XX XX      C. glutamicum derived ORF SEQ ID 1959.
XX XX

```

```

XX XX      Coryneform; nucleic acid array; fermentation; culture; ds.
XX XX
XX XX      Corynebacterium glutamicum.
XX XX
XX XX      DE10128510-A1.
XX XX
XX XX      19-DEC-2002.
XX XX
XX XX      13-JUN-2001; 2001DE-01028510.
XX XX
XX XX      13-JUN-2001; 2001DE-01028510.
XX XX
XX XX      13-JUN-2001; 2001DE-01028510.
XX XX
XX XX      (DEGS ) DEGUSA AG.
XX XX
XX XX      Farwick M, Moeckel B, Pfeifferle W, Bathe B, Hutmacher K;
XX XX      WPI, 2003-279970/28.
XX XX
XX XX      New nucleic acid array useful for monitoring mRNA expression of
XX XX      Corynebacterium glutamicum during fermentation, comprising nucleic acid
XX XX      from Corynebacterium glutamicum.
XX XX
XX XX      Claim 1; Page 647; 709pp; German.
XX XX
XX XX      This invention describes a novel nucleic acid array involving
XX XX      Corynebacterium glutamicum polynucleotides. The arrays are used to
XX XX      analyse C. glutamicum, particularly for monitoring a fermentation process
XX XX      to determine expression levels of C. glutamicum cellular mRNA. Such
XX XX      monitoring particularly differentiates between expression levels of
XX XX      different strains of C. glutamicum and allows the adjustment of different
XX XX      culture and fermentation conditions. ACA0010-ACA02188 represent C.
XX XX      glutamicum derived polynucleotides described in the disclosure of the
XX XX      invention
XX XX
XX XX      Sequence 177 BP; 54 A; 40 C; 47 G; 36 T; 0 U; 0 Other;
XX XX
XX XX      Query Match      13.9%; Score 177; DB 7; Length 177;
XX XX      Best Local Similarity 100.0%; Pred. NO. 4.8e-46;
XX XX      Matches 177; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX XX
XX XX      30 TTGCACATAGCTAGTTGAGTGGCAAACTGGCAACAAACTACCCGGCAATTGTGTG 89
XX XX      1 TTGCACATAGCTAGTTGAGTGGCAAACTGGCAACAAACTACCCGGCAATTGTGTG 60
XX XX
XX XX      90 ATGATTGATGTTGTCGCAAAAAACGCAAGATTTCAAGCTGAGAGTGTGCGCATCCA 149
XX XX      61 ATGATTGATGTTGTCGCAAAAAACGCAAGATTTCAAGCTGAGAGTGTGCGCATCCA 120
XX XX
XX XX      150 AGGAGCCCTGGAACCAAGATTAAGGTTATCGGCGCTACGAATCGCGCAAGTTC 206
XX XX      121 AGGAGCCCTGGAACCAAGATTAAGGTTATCGGCGCTACGAATCGCGCAAGTTC 177
XX XX
XX XX      RESULT 14
XX XX      ID AAH68373 standard; DNA; 711 BP.
XX XX      AAH68373;
XX XX      AAH68373;
XX XX      26-SEP-2001 (first entry)
XX XX      C glutamicum coding sequence fragment SEQ ID NO: 3408.
XX XX
XX XX      Coryneform bacterium; amino acid synthesis; vitamin; saccharide;
XX XX      organic acid synthesis; ds.
XX XX      Corynebacterium glutamicum.
XX XX
XX XX      EP1108790-A2.
XX XX      20-JUN-2001.
XX XX      18-DEC-2000; 2000EP-00127688.
XX XX

```

XX	WIPI; 2001-376931/40.	
DR	P-PSDB; AAG90036.	
XX		
PT	Novel polynucleotides derived from Corynebacterium bacteria, for identifying	
PT	mutation point of a gene, measuring expression of a gene, analyzing	
PT	expression profile or pattern of a gene and identifying homologous gene.	
XX		
PS	Claim 8; SEQ ID NO 290; 246bp + Sequence Listing; English.	
XX		
CC	The present invention provides a number of nucleotide and protein	
CC	sequences from the Corynebacterium bacterium Corynebacterium glutamicum. These	
CC	are useful for identifying the mutation point of a gene derived from a	
CC	mutant of coryneform bacterium, measuring expression amount and analysing	
CC	the expression profile or expression pattern of a gene derived from	
CC	Corynebacterium bacterium, and identifying a homologue of a gene derived from	
CC	coryneform bacterium. Corynebacterium bacteria are useful for producing amino	
CC	acids, nucleic acids, vitamins, saccharides and organic acids,	
CC	particularly L-lysine. The present sequence is a nucleic acid described	
CC	in the exemplification of the invention. Note: The sequence data for this	
CC	patent did not form part of the printed specification, but was obtained	
CC	in electronic format directly from the European Patent Office	
XX		
SQ	Sequence 324 BP; 60 A; 79 C; 80 G; 105 T; 0 U; 0 Other;	
	Query Match 25.5%; Score 324; DB 5; Length 324;	
	Best Local Similarity 100.0%; Pred. No. 4,4e-93;	
	Matches 324; Conservative 0; Mismatches 0; Indels 0; Gaps 0	
QY	853 ATGACACATGATTTCTTCCTGATTTCTCCTGTTGTGCGAGATGTGACGATTAATCTTT 912	
DB	1 ATAGCAACTGATTTCTTCCTGATTTCTCCTGTTGTGCGAGATGTGACGATTAATCTTT 60	
QY	913 GCCTCCGGCGGCGGTTCCGTTCTTAATCCCTTAAGCCCTACCTGATCACATTTG163GC 972	
DB	61 GCGCTCCGGCGGTTCCGTTCTTAATCCCTTAAGCCCTACCTGATCACATTTG163GC 120	
QY	973 AAAATGGCGATGTGATGTCGACAGAGAAATCTTGGCATTTTGACCGCATCAAGTTTCCG 1033	
DB	121 AAAATGGCGATGTGATGTCGACAGAGAAATCTTGGCATTTTGACCGCATCAAGTTTCCG 180	
QY	1033 AGCAATGCGATAGATCTGAAGACTTAACCTTTGATCTCATTTGCCGTGGCATTAACGTG 1099	
DB	181 AGCAATGCGATAGATCTGAAGACTTAACCTTTGATCTCATTTGCCGTGGCATTAACGTG 240	
QY	1093 GTGGCGCATTTCTTTGGCGGCTGACGACCTTTGTGACCGTTGGCGCTGCGACCAATCGTT 115	
DB	241 GTGGCGCATTTCTTTGGCGGCTGACGACCTTTGTGACCGTTGTGACCGTTGGCGACCAATCGTT 300	
QY	1153 TTGTGTGACTGTGATCTTTTC 1176	
DB	301 TTGTGTGACTGTGATCTTTTC 324	
	RESULT 11	
ID	ACA01970	
XX	ACA01970 standard; DNA; 324 BP.	
AC	ACA01970;	
DT	04-JUN-2003 (first entry)	
XX		
DE	C. glutamicum derived ORF SEQ ID 1961.	
XX		
KW	Coryneform; nucleic acid array; fermentation; culture; ds.	
OS	Corynebacterium glutamicum.	
XX		
PN	DE30128510-A1.	
XX		
PD	19-DEC-2002.	
XX		
FP	13-JUN-2001; 2001DE-01028510.	

Seq	Query Match	Best Local Similarity	25.5%; Score 324; DB 7; Length 324;	Matches 324; Conservative 0; Mismatches 0; Indels 0; Gaps 0
XX	13-JUN-2001; 2001DE-01028510.			
XX	(DEGS) DEGUSSA AG.			
XX	Farwick M, Moeckel B, Pfefferle W, Bathe B, Huthmacher K;			
XX	WPI; 2003-279970/28.			
XX	New nucleic acid array useful for monitoring mRNA expression of			
XX	Corynebacterium glutamicum during fermentation, comprising nucleic acid			
XX	from Corynebacterium glutamicum.			
XX	Claim 1; Page 647; 709pp; German.			
XX	This invention describes a novel nucleic acid array involving			
XX	Corynebacterium glutamicum polynucleotides. The arrays are used to			
XX	analyse C. glutamicum, particularly for monitoring a fermentation process			
XX	to determine expression levels of C. glutamicum cellular mRNA. Such			
XX	monitoring particularly differentiates between expression levels of			
XX	different strains of C. glutamicum and allows the adjustment of different			
XX	culture and fermentation conditions. ACN00010-ACN02188 represent C			
XX	glutamicum derived polynucleotides described in the disclosure of the			
XX	invention			
XX	Sequence 324 BP; 60 A; 79 C; 80 G; 105 T; 0 U; 0 Other;			
XX	Query Match	25.5%; Score 324; DB 7; Length 324;		
XX	Best Local Similarity	100.0%; Pred. No. 4,4e-93;		
XX	Matches 324; Conservative	0; Mismatches 0; Indels 0; Gaps 0		
XX	853 ATGACACTGATTTCTCTCTGATTTCTCTGTTGTGCGAGATGTGACGTCACTTAATCTTT	912		
XX	1 ATGACACTGATTTCTCTCTGATTTCTCTGTTGTGCGAGATGTGACGTCACTTAATCTTT	60		
XX	913 GCGCTCCGGGGGGTCCGTTCTTAATCCTTAAGCCCCACGTGATCAACAATTTGTGGGC	972		
XX	61 GCGCTCCGGGGGGTCCGTTCTTAATCCTTAAGCCCCACGTGATCAACAATTTGTGGGC	120		
XX	973 AAATGCGCATGTGATGCCAGCAGGAATCCTTGCCATTTTGACCGCATCAACGTTTGC	1033		
XX	121 AAATGCGCATGTGATGCCAGCAGGAATCCTTGCCATTTTGACCGCATCAACGTTTGC	180		
XX	1033 AGCAATGCGATAGATCTGAAGACTTAACCTTTGATCATTTGACCGTATTAACAGTG	1092		
XX	181 AGCAATGCGATAGATCTGAAGACTTAACCTTTGATCATTTGACCGTATTAACAGTG	240		
XX	1093 GTGGCGCATCTTTCTTGGCGGTGCAGCGACCTTTGTAAGCGTTGGCGGTGCAGCATCGTT	1155		
XX	241 GTGGCGCATCTTTCTTGGCGGTGCAGCGACCTTTGTAAGCGTTGGCGGTGCAGCATCGTT	300		
XX	1153 TTGTTGAGCTGTGATCTTTTC	1176		
XX	301 TTGTTGAGCTGTGATCTTTTC	324		
XX	RESULT 12			
XX	AAFe1688/C			
XX	ID AAF61688 standard; DNA; 715 BP.			
XX	AAFe1688;			
XX	DT 12-UTU-2001 (first entry)			
XX	C. glutamicum lrp encoding DNA.			
XX	Fermentation; L-amino acid production; lrp gene; lysine; isoleucine;			
XX	Medicine; animal feed supplement; ds.			
XX	Corynebacterium glutamicum.			
XX	Key	Location/Qualifiers		
XX	-35_signal	62..67		

QY 769 TCACAGTCAGGCGCTATTGCGCGCTGCTGATCTCTTCTTGAGGTCTGTTGACCATCCGTA 828
 DB 480 TCACAGTCAGGCGCTATTGCGCGCTGCTGATCTCTTCTTGAGGTCTGTTGACCATCCGTA 539
 QY 829 CTTCTCTTTGGGAAAAGGCTGCTTAA 853
 DB 540 CTTCTCTTTGGGAAAAGGCTGCTTAA 564

RESULT 7

AAAF61693/C
 ID AAF61693 standard; DNA; 397 BP.

AAAF61693;
 AC AAF61693;
 DT 12-JUL-2001 (first entry)

DE C. glutamicum lrp encoding DNA fragment SEQ ID 7.

XX Fermentation; L-amino acid production; lrp gene; lysine; isoleucine;
 KM medicine; animal feed supplement; ds.
 XX Corynebacterium glutamicum.

OS EPI090993-A1.

PN 11-APR-2001.

PF 29-SEP-2000; 2000EP-00121159.

PR 05-OCT-1999; 99DE-01047792.

PA (DEBS) DEGUSA-HUELS AG.

PI Moeckel B, Pfeifferle W, Puehler A, Kalinowski J, Bathe B;

DR WPI; 2001-292927/31.

PT New lrp gene from coryneform bacteria, used to prepare transformants with
 PT increased synthesis of amino acids, particularly lysine and isoleucine.

PS Example 3; Page 16; 22pp; German.

XX This invention describes a novel isolated nucleic acid (I) from
 CC coryneform bacteria used for the fermentative production of selected L-
 CC amino acids, by fermenting the amino acid-producing coryneform in which
 CC at least the lrp gene has been weakened or amplified, then isolating
 CC amino acids that have accumulated in the medium or cells. (I) is used to
 CC transform coryneforms for production of L-amino acids, specifically
 CC lysine and isoleucine, which are used in medicine and particularly as
 CC animal feed supplement. It may also be used as probes and primers for
 CC isolating related sequences. Regulating expression of (I) improves
 CC production of amino acids, especially of L-lysine. This sequence encodes
 CC a fragment of the Corynebacterium glutamicum lrp protein which is used in
 CC the method described in the invention

CC Sequence 397 BP; 96 A; 100 C; 104 G; 97 T; 0 U; 0 Other;

QY Query Match 28.8%; Score 365.6; DB 4; Length 397;

DB Best Local Similarity 98.9%; Pred. No. 2.2e-106;

Matches 368; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 9 AATGGAATCTAGCTCATATATTGCAACATAGCTTAGTGGCTGCGAAATGGCAACA 68

DB 376 AATGGAATCTAGCTCATATATTGCAACATAGCTTAGTGGCTGCGAAATGGCAACA 317

QY 69 AAACATCCCGGCAATGTTGATGATGTTAGTGTGCAAAAAAGCAAGAGATTCAATCA 128

DB 316 AAACATCCCGGCAATGTTGATGATGTTAGTGTGCAAAAAAGCAAGAGATTCAATCA 257

QY 129 GCCTGAGGTGTGCGCATCCAAAGCAGCCCTGGAACAGATATTAAGTTATCGGCGCT 188

DB 256 GCCTGAGGTGTGCGCATCCAAAGCAGCCCTGGAACAGATATTAAGTTATCGGCGCT 197
 QY 189 ACGAAATCCGCAAGGCTTAAAAACCTCCCTTGCTGCAAGTTGGCATGTACCGATTG 248
 DB 196 ACGAAATCCGCAAGGCTTAAAAACCTCCCTTGCTGCAAGTTGGCATGTACCGATTG 137
 QY 249 GTATTCGCTTTGGTCTCTTGGTTATTCAATAGCGCTACGAATGGTGGAGCCCACTGT 308
 DB 136 GTATTCGCTTTGGTCTCTTGGTTATTCAATAGCGCTACGAATGGTGGAGCCCACTGT 77
 QY 309 TTTCGAGCTGATTTTCGCGGCTCCACCGAAATGCTGTCAATCGCCCTGTTGTGGCG 368
 DB 76 TTTCGAGCTGATTTTCGCGGCTCCACCGAAATGCTGTCAATCGCCCTGTTGTGGCG 17
 QY 369 CAGCGCCCTGG 380
 DB 16 CAGCGCCCTGG 5

RESULT 8

AAAF61695/C
 ID AAF61695 standard; DNA; 778 BP.

AAAF61695;
 AC AAF61695;
 DT 12-JUL-2001 (first entry)

DE C. glutamicum lrp encoding DNA fragment SEQ ID 9.

XX Fermentation; L-amino acid production; lrp gene; lysine; isoleucine;
 KM medicine; animal feed supplement; ds.
 XX Corynebacterium glutamicum.

OS EPI090993-A1.

PN 11-APR-2001.

PF 29-SEP-2000; 2000EP-00121159.

PR 05-OCT-1999; 99DE-01047792.

PA (DEBS) DEGUSA-HUELS AG.

PI Moeckel B, Pfeifferle W, Puehler A, Kalinowski J, Bathe B;

DR WPI; 2001-292927/31.

PT New lrp gene from coryneform bacteria, used to prepare transformants with
 PT increased synthesis of amino acids, particularly lysine and isoleucine.

PS Example 3; Page 17; 22pp; German.

XX This invention describes a novel isolated nucleic acid (I) from
 CC coryneform bacteria used for the fermentative production of selected L-
 CC amino acids, by fermenting the amino acid-producing coryneform in which
 CC at least the lrp gene has been weakened or amplified, then isolating
 CC amino acids that have accumulated in the medium or cells. (I) is used to
 CC transform coryneforms for production of L-amino acids, specifically
 CC lysine and isoleucine, which are used in medicine and particularly as
 CC animal feed supplement. It may also be used as probes and primers for
 CC isolating related sequences. Regulating expression of (I) improves
 CC production of amino acids, especially of L-lysine. This sequence encodes
 CC a fragment of the Corynebacterium glutamicum lrp protein which is used in
 CC the method described in the invention

CC Sequence 778 BP; 175 A; 208 C; 207 G; 188 T; 0 U; 0 Other;

QY Query Match 28.8%; Score 365.6; DB 4; Length 778;

DB Best Local Similarity 98.9%; Pred. No. 3.2e-106;

Matches 368; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 9 AATGGAATCTAGCTCATATATTGCAACATAGCTTAGTGGCTGCGAAATGGCAACA 68

CC acids, nucleic acids, vitamins, saccharides and organic acids,
 particularly L-lysine. The present sequence is a nucleic acid described
 in the exemplification of the invention. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 in electronic format directly from the European Patent Office

XX Sequence 753 BP; 140 A; 211 C; 194 G; 208 T; 0 U; 0 Other;

Query Match 59.2%; Score 753; DB 5; Length 753;

Best Local Similarity 100.0%; Pred. No. 1.3e-230; Indels 0; Gaps 0;

Matches 753; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 101 GTGCAAAAACGCAAGATTCATTCAGACCTGAGAGTGTGCCATCCAGAGCCCTG 160
 DB 1 GTGCAAAAACGCAAGATTCATTCAGACCTGAGAGTGTGCCATCCAGAGCCCTG 60
 QY 161 GAACCAATGATTAAGGTTATCGCGCTACGAAATCGGCGCAAGGCTTAAACCTCCCT 220
 DB 61 GAACCAATGATTAAGGTTATCGCGCTACGAAATCGGCGCAAGGCTTAAACCTCCCT 120
 QY 221 GTGCAAGTTGGGCGATGACCGATGATGATGATGATGATGATGATGATGATGATGAT 280
 DB 121 GTGCAAGTTGGGCGATGACCGATGATGATGATGATGATGATGATGATGATGATGAT 180
 QY 281 GGCTACGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 340
 DB 181 GGCTACGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 240
 QY 341 ATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 400
 DB 241 ATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300
 QY 401 TTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 460
 DB 301 TTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 360
 QY 461 AACCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 520
 DB 361 AACCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420
 QY 521 GCGGCGAGGCGCGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 580
 DB 421 GCGGCGAGGCGCGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
 QY 581 TCTTACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 640
 DB 481 TCTTACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 540
 QY 641 ATTAAGGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 700
 DB 541 ATTAAGGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
 QY 701 CGAACGAAAAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 760
 DB 601 CGAACGAAAAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
 QY 761 GTGGTAATTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 820
 DB 661 GTGGTAATTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
 QY 821 ATCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 883
 DB 721 ATCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 783

RESULT 6

ACAO1969 standard; DNA; 564 BP.

XX ACAA01969;
 XX AC
 XX ID
 XX DT
 XX 04-JUN-2003 (first entry)

DE C. glutamicum derived ORF SEQ ID 1960.

XX Corynebacterium glutamicum; fermentation; culture; de.

XX Corynebacterium glutamicum.

XX DE10128510-A1.

XX 19-DEC-2002.

XX 13-JUN-2001; 2001DE-01028510.

XX 13-JUN-2001; 2001DE-01028510.

XX (DEGS) DEGUSA AG.

XX Farwick M, Moeckel B, Pfefferle W, Bathe B, Huthmacher K;

XX WPI; 2003-279970/28.

PT New nucleic acid array useful for monitoring mRNA expression of
 PT Corynebacterium glutamicum during fermentation, comprising nucleic acid
 PT from Corynebacterium glutamicum.

XX Claim 1; Page 647; 709pp; German.

CC This invention describes a novel nucleic acid array involving
 CC Corynebacterium glutamicum polynucleotides. The arrays are used to
 CC analyze C. glutamicum, particularly for monitoring a fermentation process
 CC to determine expression levels of C. glutamicum cellular mRNA. Such
 CC monitoring particularly differentiates and allows the adjustment of different
 CC culture and fermentation conditions. ACAA0010-ACAA0218 represent C.
 CC glutamicum derived polynucleotides described in the disclosure of the
 CC invention

XX Sequence 564 BP; 91 A; 167 C; 144 G; 162 T; 0 U; 0 Other;

Query Match 43.5%; Score 553; DB 7; Length 564;

Best Local Similarity 99.8%; Pred. No. 1.8e-166; Indels 1; Gaps 1;

Matches 564; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 289 ATGTGGGAGAGCCCACTGTTTCCGCTGATTTTCCGCGCTCCACCAATGCTGCT 348
 DB 1 ATGTGGGAGAGCCCACTGTTTCCGCTGATTTTCCGCGCTCCACCAATGCTGCT 60
 QY 349 CATGCCCTGTTGTGGGCGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 408
 DB 61 CATGCCCTGTTGTGGGCGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 119
 QY 409 GAATTCGCGCAGATTTCTATGCGTTTCAATTCGCGCTGATGATGATGATGATGATGAT 468
 DB 120 GAATTCGCGCAGATTTCTATGCGTTTCAATTCGCGCTGATGATGATGATGATGATGAT 179
 QY 469 TGGCGGTTTCTATGCGTTTGTGGGCTTATGCGAGAGGCTTACGATGATGATGATGATGAT 528
 DB 180 TGGCGGTTTCTATGCGTTTGTGGGCTTATGCGAGAGGCTTACGATGATGATGATGATGAT 239
 QY 529 GCCCGAGGCTGATGCGGCTGATGCGGCTTATGCGGCTTATGCGGCTTATGCGGCTTAT 588
 DB 240 GCCCGAGGCTGATGCGGCTGATGCGGCTTATGCGGCTTATGCGGCTTATGCGGCTTAT 299
 QY 589 GGTATTCGCGGCTGATGCGGCTGATGCGGCTTATGCGGCTTATGCGGCTTATGCGGCTTAT 648
 DB 300 GGTATTCGCGGCTGATGCGGCTGATGCGGCTTATGCGGCTTATGCGGCTTATGCGGCTTAT 359
 QY 649 CCTGAGTTGCGGCTTATGCGGCTTATGCGGCTTATGCGGCTTATGCGGCTTATGCGGCTTAT 708
 DB 360 CCTGAGTTGCGGCTTATGCGGCTTATGCGGCTTATGCGGCTTATGCGGCTTATGCGGCTTAT 419
 QY 709 AAAGCAGATCCCTTCTGCTGCTGCTGCGAGGTTTAAAGCTTACCAATGCTCTTGCTGAT 768
 DB 420 AAAGCAGATCCCTTCTGCTGCTGCTGCGAGGTTTAAAGCTTACCAATGCTCTTGCTGAT 479

XX 02-MAY-2001.
 PD 11-OCT-2000; 2000EP-00122057.
 XX 27-OCT-1999; 99DE-01051708.
 PR (DEGS) DEGUSSA AG.
 PA (KERU) FORSCHUNGSZENTRUM JUELICH GMBH.
 PI Kernerkecht N, Eggeling L, Sahn H, Pfeifferle W;
 XX WPI: 2001-391595/42.
 DR P-PSDB; AAB66247.
 PT New export genes from coryneform bacteria, useful for increasing
 PT fermentative production of branched-chain amino acids.
 XX Claim 5; Page 14-15; 23pp; German.
 PS
 CC This invention describes a novel isolated polynucleotide (1) containing
 CC at least one sequence that (i) is 70% identical with a sequence that
 CC encodes at least one of 251 amino acid (aa) (3) or 108 aa (5)
 CC polypeptides; (11) encodes a polypeptide at least 70% identical with (3)
 CC or (5); (11) is the complement of (1) or (11), or (11) contains at least
 CC 15 consecutive bases from (1) - (11). The invention also describes (a)
 CC protein derived from sequences of 753 bp (2) or 324 bp (4); (b)
 CC coryneform microorganisms, especially Corynebacterium, transformed with
 CC one or more (1), where these are replicative DNA; (c) production of
 CC branched-chain L-aa by fermentation of coryneform bacteria in which the
 CC brn and/or brn genes (or equivalent sequences) are amplified
 CC especially overexpressed; and (d) method for isolating the brn and/or
 CC brn genes. (1) is used for transformation of coryneform bacteria being
 CC used for fermentative production of branched-chain amino acids,
 CC specifically leucine, isoleucine and valine, which are useful in medicine
 CC and animal nutrition. (1) can also be used as source of primers and
 CC probes for isolation of related sequences. Transformation with (1)
 CC increases yield of branched-chain amino acids. This sequence encodes the
 CC Corynebacterium glutamicum ATCC 14752 brn protein described in the
 CC method of the invention
 CC
 CC Sequence 753 BP; 140 A; 211 C; 194 G; 208 T; 0 U; 0 Other;
 SQ
 Query Match 59.2%; Score 753; DB 4; Length 753;
 Best Local Similarity 100.0%; Pred. No. 1.3e-230; Indels 0; Gaps 0;
 Matches 753; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 101 GTGCAAAAACGCAAGATTCATTCAGACCTGAGGTGTGCGCATCCAGGACGCCCTG 160
 DB 1 GTGCAAAAACGCAAGATTCATTCAGACCTGAGGTGTGCGCATCCAGGACGCCCTG 60
 QY 161 GAACCAATGATTAAGTTATGCGGCTTACGAAATCGGGCAAGCTCTAAAACCTCCCTT 220
 DB 61 GAACCAATGATTAAGTTATGCGGCTTACGAAATCGGGCAAGCTCTAAAACCTCCCTT 120
 QY 221 GTGCAAGTTGGGCAATGATACCGCATTTGGTATTTGGTCTCTTGGTATTCATATC 280
 DB 121 GTGCAAGTTGGGCAATGATACCGCATTTGGTATTTGGTCTCTTGGTATTCATATC 180
 QY 281 GGCTACGAATGTTGGGCAAGCTCTGTTTCCGGCTGATTTTCCGGGCTTCCACCGAA 340
 DB 181 GGCTACGAATGTTGGGCAAGCTCTGTTTCCGGCTGATTTTCCGGGCTTCCACCGAA 240
 QY 341 ATGCTGTGATGCGCCTGTTGGGCGAGGCGCCCTGCGGCATCGCGTCAACCA 400
 DB 241 ATGCTGTGATGCGCCTGTTGGGCGAGGCGCCCTGCGGCATCGCGTCAACCA 300
 QY 401 TTGCTGTGAATCTTCGCAAGTATTCATTCATTCATTCATTCATTCATTCATTCATTC 460
 DB 301 TTGCTGTGAATCTTCGCAAGTATTCATTCATTCATTCATTCATTCATTCATTCATTC 360
 QY 461 AACCCATGCGCGTTTCTATTCGTTTCCGGCTTATGAGGAGGCTTACGCAATCACT 520

DB 361 AACCCATGCGCGTTTCTATTCGTTTCCGGCTTATGAGGAGGCTTACGCAATCACT 420
 QY 521 GCGGCAAGCGCCGCAAGGTGTGCGGCTGAGGCAATTCATTCATTCATTCATTCATTC 580
 DB 421 GCGGCAAGCGCCGCAAGGTGTGCGGCTGAGGCAATTCATTCATTCATTCATTCATTC 480
 QY 581 TCTTACTGGTATTCGCGGCTGTCTACCGGAGTGGCGATTCGAGGATTCCTTTGAA 640
 DB 481 TCTTACTGGTATTCGCGGCTGTCTACCGGAGTGGCGATTCGAGGATTCCTTTGAA 540
 QY 641 ATTAAGGCGCTGAGTTCGCGCCTTGTCTCTCTCTTGTCAAGCTGACTTGGATTCCTGC 700
 DB 541 ATTAAGGCGCTGAGTTCGCGCCTTGTCTCTCTCTTGTCAAGCTGACTTGGATTCCTGC 600
 QY 701 CGAACGAAAAGACAGATTCCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 760
 DB 601 CGAACGAAAAGACAGATTCCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
 QY 761 GTGTAATTCAGATTCAGGCGCTTATTTGGGCGCTGCTGATCTTGTGGTCTGTGACC 820
 DB 661 GTGTAATTCAGATTCAGGCGCTTATTTGGGCGCTGCTGATCTTGTGGTCTGTGACC 720
 QY 821 ATCCGTAATTCCTTCTTGGGAAAGCTGTCTAA 853
 DB 721 ATCCGTAATTCCTTCTTGGGAAAGCTGTCTAA 753
 RESULT 5
 AAH65254
 ID AAH65254 standard; DNA; 753 BP.
 XX
 AC AAH65254;
 DT 26-SEP-2001 (first entry)
 XX
 DE C glutamicum coding sequence fragment SEQ ID NO: 289.
 XX
 KM Coryneform bacterium; amino acid synthesis; vitamin; saccharide;
 XX organic acid synthesis; de.
 XX
 OS Corynebacterium glutamicum.
 XX
 PN EPI108790-A2.
 XX
 PD 20-JUN-2001.
 XX
 PF 18-DEC-2000; 2000EP-00127688.
 XX
 XX 16-DEC-1999; 99DP-00377484.
 PR 07-APR-2000; 2000JP-00159162.
 PR 03-AUG-2000; 2000JP-00280988.
 XX
 PA (KYOW) KYOWA HAKKO KOGYO KK.
 XX
 XX Nakagawa S, Mizoguchi H, Ardo S, Hayashi M, Ochiai K, Yokoi H;
 PT Tateishi N, Senoh A, Ikeda M, Ozaki A;
 DR WPI: 2001-376931/40.
 DR P-PSDB; AAG90035.
 XX
 PT Novel polynucleotides derived from Coryneform bacteria, for identifying
 PT mutation point of a gene, measuring expression of a gene, analyzing
 PT expression profile or pattern of a gene and identifying homologous gene.
 XX
 PS Claim 8; SEQ ID NO 289; 246pp + Sequence Listing; English.
 XX
 CC The present invention provides a number of nucleotide and protein
 CC sequences from the Coryneform bacterium Corynebacterium glutamicum. These
 CC are useful for identifying the mutation point of a gene derived from a
 CC mutant of coryneform bacterium, measuring expression amount and analysing
 CC the expression profile or expression pattern of a gene derived from
 CC Coryneform bacterium, and identifying a homologue of a gene derived from
 CC coryneform bacterium. Coryneform bacteria are useful for producing amino

Query Match	Best Local Similarity	100.0%	Score 1271	DB 5	Length 349980	Matches 1271	Conservative	100.0%	0	Mismatches	0	Indels	0	Gaps	0
PD	20-JUN-2001.														
PP	18-DEC-2000; 2000EP-00127688.														
PR	16-DEC-1999; 98UP-00377484.														
PR	07-APR-2000; 2000JP-00159152.														
PR	03-AUG-2000; 2000JP-00280988.														
XX	(KYOW) KYOWA HAKKO KOXOYO KK.														
PA	Nakagawa S, Mizuguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;														
PI	Tateishi N, Senoh A, Ikeda M, Ozaki A,														
DR	WPI; 2001-376931/40.														
PT	Novel polynucleotides derived from Corynebacterium bacteria, for identifying														
PT	mutation point of a gene, measuring expression of a gene, analyzing														
PT	expression profile or pattern of a gene and identifying homologous gene.														
XX	Claim 7; SEQ ID NO 1; 246bp + Sequence Listing; English.														
XX	The present invention provides a number of nucleotide and protein														
CC	sequences from the Corynebacterium bacterium Corynebacterium glutamicum. These														
CC	are useful for identifying the mutation point of a gene derived from a														
CC	mutant of corynebacterium bacterium, measuring expression amount and analysing														
CC	the expression profile or expression pattern of a gene derived from														
CC	Corynebacterium bacterium, and identifying a homologue of a gene derived from														
CC	Corynebacterium bacterium. Corynebacterium bacteria are useful for producing amino														
CC	acids, nucleic acids, vitamins, saccharides and organic acids,														
CC	particularly L-lysine. The present sequence is a nucleic acid described														
CC	in the exemplification of the invention. Note: The sequence data for this														
CC	patent did not form part of the printed specification, but was obtained														
CC	in electronic format directly from the European Patent Office														
SEQ	Sequence 349980 BP; 79703 A; 91547 C; 98381 G; 80349 T; 0 U; 0 Other;														
Query Match	Best Local Similarity	100.0%	Score 1271	DB 5	Length 349980	Matches 1271	Conservative	100.0%	0	Mismatches	0	Indels	0	Gaps	0
QY	276729	GCGGCATCAATGCAATCTAGCTTCAATATATGACCAATAGCCTAGTGGAGGTGGCAAC	60												
DB	276729	GCGGCATCAATGCAATCTAGCTTCAATATATGACCAATAGCCTAGTGGAGGTGGCAAC	276788												
QY	61	TGGCAACAAACTACCCCGCAATTGTGTGAGCATTTGATAGTGTGCAAAAACGCAAGAT	120												
DB	276789	TGGCAACAAACTACCCCGCAATTGTGTGAGCATTTGATAGTGTGCAAAAACGCAAGAT	276848												
QY	121	TCATTCAGGCTGAGAGGTGTGCGCATCCAGGCAAGCCCTGGAAACAGATGATTAAGTTA	180												
DB	276849	TCATTCAGGCTGAGAGGTGTGCGCATCCAGGCAAGCCCTGGAAACAGATGATTAAGTTA	276908												
QY	181	TGGGCGCAAGCAATCCGGCAAGGCTTAAAAACCTCCCTTGCGCAGAGTTGGGCAATGA	240												
DB	276909	TGGGCGCAAGCAATCCGGCAAGGCTTAAAAACCTCCCTTGCGCAGAGTTGGGCAATGA	276968												
QY	241	CCGATATGATATGCGGTTGGTCTCTTGGTATTCAATAGCGCATAGCAATGATGGGAGC	300												
DB	276969	CCGATATGATATGCGGTTGGTCTCTTGGTATTCAATAGCGCATAGCAATGATGGGAGC	277028												
QY	301	CCGATATGATATGCGGTTGGTCTCTTGGTATTCAATAGCGCATAGCAATGATGGGAGC	360												
DB	277029	CCGATATGATATGCGGTTGGTCTCTTGGTATTCAATAGCGCATAGCAATGATGGGAGC	277088												
QY	361	TGTGGGCGAGCGCCCTGGGCGCCATGCGCTACACCAATGCTGTGTAACCTTCGCA													

D	b		277209	TTCGGTTTCCGCGCTTAATGCAGAAAGCCTAAGCAATCACTACGCCGACCAGCCCCGACGGCTG	277258
O	y		541	GTCGGCGGTGGCGACTTATCTCATATGCATAATAGCGTTTCACTCTACTGGGTATTCCGCGG	600
D	b		277269	GTCGGGGTGCGACAATTATCTCATATGCATAATAGCGTTTCACTCTCACTACGGGTATTCCGCGG	277328
O	y		601	TCTCACCGAGATGCGCATGCGAGAGTTGATTCCTTTTGAATTTAAGGGCTCGAGTTCCG	660
D	b		277329	TCTCACCGAGATGCGCATGCGAGAGTTGATTCCTTTTGAATTTAAGGGCTCGAGTTCCG	277388
O	y		661	CCTTTGCTCTCTCTCTTTTGTACGCGCTGACTTTGGATTCCGTCCGCAACGAAAAAGCATCCC	720
D	b		277389	CCCTTCTCTCTCTCTTTGTACGCGCTGACTTTGGATTCCGTCCGCAACGAAAAAGCATATCCC	277448
O	y		721	TTCCTGCTGCTCGCAGAGTTTAGCCTTACCATTTGCTCTTGGTGTAATTCACAGTACAGC	780
D	b		277449	TTCCTGCTGCTCGCAGAGTTTAGCCTTACCATTTGCTCTTGGTGTAATTCACAGTACAGC	277508
O	y		781	CCGATTTGGCGGGCGCTGCTGATCTTTTGGGAGTCTGTGACCATCCGGTACTCTTCTTGGG	840
D	b		277509	CCGATTTGGCGGGCGCTGCTGATCTTTTGGGAGTCTGTGACCATCCGGTACTCTTCTTGGG	277568
O	y		841	AAGAGCTGCTMAATGACAACTGATTTCTCTGTAATTCCTCTGTGTTCGAGATATGCA	900
D	b		277569	AAGAGCTGCTMAATGACAACTGATTTCTCTGTAATTCCTCTGTGTTCGAGATATGCA	277628
O	y		901	GTCATTACTTTTGGCGCTCCGGCGGGTCCGTTCTTAATCCTTAAGCCCTTACGTGAATCA	960
D	b		277629	GTCATTACTTTTGGCGCTCCGGCGGGTCCGTTCTTAATCCTTAAGCCCTTACGTGAATCA	277688
O	y		961	CAATTGTGTGGGCAAAATGGCGCATGTGATGCGCACAGAGAAATCCTTGCCATTTTGACCGCA	1020
D	b		277689	CAATTGTGTGGGCAAAATGGCGCATGTGATGCGCACAGAGAAATCCTTGCCATTTTGACCGCA	277748
O	y		1021	TCAAAGCTTTCCGACGAATGCCATTAATCTGAAGAATCTAACCTTTGGTCTCATTTGGCGCTT	1080
D	b		277749	TCAAAGCTTTCCGACGAATGCCATTAATCTGAAGAATCTAACCTTTGGTCTCATTTGGCGCTT	277808
O	y		1081	GCGATTACAGTGTGGCGCATCTTCTTGGCGGTGACGACACTTTGTGAGCGTTGCGCT	1140
D	b		277809	GCGATTACAGTGTGGCGCATCTTCTTGGCGGTGACGACACTTTGTGAGCGTTGCGCT	277868
O	y		1141	GGAGACATGCTTTTTGTGTGACCTGGTGAATCTTTTCTPAAAATGCGCAATAAATACAAAAAT	1200
D	b		277869	GGAGACATGCTTTTTGTGTGACCTGGTGAATCTTTTCTPAAAATGCGCAATAAATACAAAAAT	277928
O	y		1201	CGCGATGCGCCTCAATTGTGAAGGGAGTGGCGATTTTTPAAGAACTTAAGAAAGCCTTAAG	1260
D	b		277929	CGCGATGCGCCTCAATTGTGAAGGGAGTGGCGATTTTTPAAGAACTTAAGAAAGCCTTAAG	277988
O	y		1261	CAGACAGCGCT 1271	
D	b		277989	CAGACAGCGCT 277999	
<hr/>					
RESULT 4					
AAH21110					
ID	AAH21110 standard; DNA; 753 BP.				
XX	AAH21110;				
XX	AAH21110;				
DT	05-SEP-2001 (first entry)				
XX	C. glutamicum brnf DNA.				
DE					
XX					
KV	L-amino acid production; brnF, brnE; branched-chain amino acid;				
KW	corynebacterium bacterium; leucine; isoleucine; valine; medicine;				
KX	animal nutrition; ds.				
OS	Corynebacterium glutamicum.				
XX					
FN	EPI096010-AI.				

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using SW model

Run on: April 19, 2004, 10:56:53 ; Search time 591.9 Seconds
(without alignments)
9122.247 Million cell updates/sec

Title: US-10-608-504-6
Perfect score: 1271
Sequence: 1 ggcgcgcataacggaatcctag.....agccttaagcagacagcgc 1271

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 212409041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_23Jan04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002s:*
7: geneseqn2003as:*
8: geneseqn2003bs:*
9: geneseqn2003cs:*
10: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1271	100.0	1271	AAH21109	Aah21109 C. glutam
2	1271	100.0	1271	AAH21112	Aah21112 C. glutam
3	1271	100.0	349980	AAH64966	Aah64966 C. glutam
4	753	59.2	753	AAH21110	Aah21110 C. glutam
5	753	59.2	753	AAH65254	Aah65254 C. glutam
6	553	43.5	564	AAH61969	Aah61969 C. glutam
7	365.6	28.8	397	AAH61693	Aah61693 C. glutam
8	365.6	28.8	778	AAH61695	Aah61695 C. glutam
9	324	25.5	324	AAH21111	Aah21111 C. glutam
10	324	25.5	324	AAH65255	Aah65255 C. glutam
11	324	25.5	324	AAH61970	Aah61970 C. glutam
12	185	14.6	715	AAH61688	Aah61688 C. glutam
13	177	13.9	177	AAH61968	Aah61968 C. glutam
14	127.4	10.0	711	AAH68373	Aah68373 C. glutam
15	127.4	10.0	711	AAH61294	Aah61294 C. glutam
16	127.4	10.0	309400	AAH68334	Aah68334 C. glutam
17	48.2	3.8	2000	ADAV1938	Adav1938 Rice gene
18	48	3.8	10732	AAH10594	Aah10594 Gene enco
19	40.8	3.2	7645	AAH59610	Ades9610 Human gen
20	40.2	3.2	4947	AAH69227	Aah69227 Tsp1.7 po
21	39.8	3.1	768	ADCG1639	Adcg1639 E. faeciu
22	39.2	3.1	6741	AAH10595	Aah10595 Gene enco
23	38.2	3.0	946	AAH5683	Aah5683 Human bre

ALIGNMENTS

```

RESULT 1
AAH21109
ID AAH21109 standard; DNA; 1271 BP.
XX
XX AAH21109;
XX
XX 05-SEP-2001 (first entry)
XX
XX C. glutamicum DNA encoding brnF and brnE.
XX
XX L-amino acid production; brnF; brnE; branched-chain amino acid;
XX Corynebacterium; leucine; isoleucine; valine; medicine;
XX animal nutrition; ds.
XX
XX OS Corynebacterium glutamicum.
XX
XX Key Location/Qualifiers
XX CDS 101..856
XX FT /*tag= a "brnF"
XX FT /product= "brnF"
XX FT 853..1179
XX FT /*tag= b
XX FT /product= "brnE"
XX
XX EP1096010-A1.
XX
XX 02-MAY-2001.
XX
XX 11-OCT-2000; 2000EP-00122057.
XX
XX 27-OCT-1999; 99DE-01051708.
XX
XX (DEGS ) DEGUSSA AG.
XX (KERU ) FORSCHUNGSZENTRUM JUELICH GMBH.
XX
XX Kemerknecht N, Eggeling L, Sahn H, Pfeifferle W,
XX MPI, 2001-391595/42.
XX P-PSDB; AAB86247, AAB86248.
XX
XX New export genes from coryneform bacteria, useful for increasing
XX fermentative production of branched-chain amino acids.
XX Claim 4 (1); Page 13; 23pp; German.
XX
XX This invention describes a novel isolated polynucleotide (I) containing

```

```

Db      181 GGGTACGATGGTGGGACAGCCCACTGTTTCGGGCTGATTTGGCGGGCTCCACCGAA 240
Qy      341 ATGCTGTCATGAGCCCTGCTGTGTGGGCGAGAGCCCTGGGGCCATGGCGCTCACCA 400
Db      241 ATGCTGTCATGAGCCCTGCTGTGTGGGCGAGAGCCCTGGGGCCATGGCGCTCACCA 300
Qy      401 TTGCTGATGACTTCCGCCACGATTTCTATGCGTTTCAATCCCGTCATGGTCAAA 460
Db      301 TTGCTGATGACTTCCGCCACGATTTCTATGCGTTTCAATCCCGTCATGGTCAAA 360
Qy      461 AACCCATGCCCCGTTTCTATTCGCTTTGCGGCTTATGACGAGAGCTTACGACT 520
Db      361 AACCCATGCCCCGTTTCTATTCGCTTTGCGGCTTATGACGAGAGCTTACGACT 420
Qy      521 GCGGCGAGGCGCGAGGCTGTGGGCGTGGCGACTTATCTCAATGCAATAGCGTTTAC 580
Db      421 GCGGCGAGGCGCGAGGCTGTGGGCGTGGCGACTTATCTCAATGCAATAGCGTTTAC 480
Qy      581 TCCTACTGGTATTCGCGGCTCTCACCGAGTGGCGATGCGAGATTGATTCCTTTGAA 640
Db      481 TCCTACTGGTATTCGCGGCTCTCACCGAGTGGCGATGCGAGATTGATTCCTTTGAA 540
Qy      641 ATTAAGGCGCTCGAGTTCGCGCTTGTCTCTCTTTGTGACGCTGACTTTGATTCG 700
Db      541 ATTAAGGCGCTCGAGTTCGCGCTTGTCTCTCTTTGTGACGCTGACTTTGATTCG 600
Qy      701 CGAAGCAAAAGAGATCCCTCTCTGCTGCTGCAAGTTTGAAGCTTACCATGCTTT 760
Db      601 CGAAGCAAAAGAGATCCCTCTCTGCTGCTGCAAGTTTGAAGCTTACCATGCTTT 660
Qy      761 GTGTATTTTCAGGTCAGGCGCTATTTGGCGGCTGTGATCTTCTGGGTCGTGACC 820
Db      661 GTGTATTTTCAGGTCAGGCGCTATTTGGCGGCTGTGATCTTCTGGGTCGTGACC 720
Qy      821 ATCCGACTTCTCTTGGGAAAGGCTGTAA 853
Db      721 ATCCGACTTCTCTTGGGAAAGGCTGTAA 753

```

RESULT 15
 AX137083/c 397 bp DNA linear PAT 30-MAY-2001
 LOCUS AX137083
 DEFINITION Sequence 7 from Patent EP1090933.
 ACCESSION AX137083
 VERSION AX137083.1 GI:14273428
 KEYWORDS
 SOURCE
 ORGANISM
 Corynebacterium glutamicum
 Corynebacterium glutamicum
 Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 Corynebacterineae; Corynebacteriaceae; Corynebacterium.

REFERENCE
 1 Mosckel,B., Pfeifferle,W., Puehler,A., Kalinowski,J. and Bathe,B.
 Nucleotide sequences coding for the lrp gene
 Patent: EP 1090933-A 7 11-Apr-2001;
 Degussa AG (DE)

FEATURES
 source
 1..397
 location/Qualifiers
 /organism="Corynebacterium glutamicum"
 /mol_type="unassigned DNA"
 /db_xref="taxon:1718"
 /note="Lrp-Tail 1"

ORIGIN

Query Match 28.8%; Score 365.6; DB 6; Length 397;
 Best Local Similarity 98.9%; Pred.No.2e-89;
 Matches 368; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```

Qy      9 AATGGAATCTAGCTTCATATTTTGACAATAAGCCCTAGTTGAGTGCGCAACTGGCAACA 68
Db      376 AATGGAATCTAGCTTCATATTTTGACAATAAGCCCTAGTTGAGTGCGCAACTGGCAACA 317
Qy      69 AAACACCGGCAATTGTGTATGATGTAGTGTGCAAAAAACGCAAGAGATTCATTCPA 128

```

```

Db      316 AAACACCGGCAATTGTGTATGATGTAGTGTGCAAAAAACGCAAGAGATTCATTCPA 257
Qy      129 GCTGAGAGTGTGCGCATCCAGGACAGCCCTGGAAACAGATGATTAAGTTATCGCGCT 188
Db      256 GCTGAGAGTGTGCGCATCCAGGACAGCCCTGGAAACAGATGATTAAGTTATCGCGCT 197
Qy      189 ACGAATGCGCGCAAGGTCTAATAAACTCCTTGTGCTGAGGTTGGGCAATGACCATTTG 248
Db      196 ACGAATGCGCGCAAGGTCTAATAAACTCCTTGTGCTGAGGTTGGGCAATGACCATTTG 137
Qy      249 GTATTGGTTGGTCTCTGTTGTTATTCATACGAGTACGAATGATGAGGCGACGCCACTGT 308
Db      136 GTATTGGTTGGTCTCTGTTGTTATTCATACGAGTACGAATGATGAGGCGACGCCACTGT 77
Qy      309 TTTCCGCGCTGATTTTTCGCGGCTCCACCGAAATGCTGTGATGCGCTTGTGTGCGG 368
Db      76 TTTCCGCGCTGATTTTTCGCGGCTCCACCGAAATGCTGTGATGCGCTTGTGTGCGG 17
Qy      369 CAGCGCCCTGCG 380
Db      16 CAGCGCCCTGCG 5

```

Search completed: April 19, 2004, 18:30:05
 Job time : 5391 secs

```

REFERENCE
AUTHORS      1 (bases 1 to 753)
TITLE        Kernerukunehito, N., Sahn, H., Eggering, L. and Pfeifferle, W.
JOURNAL      Nucleotide sequence encoding the discharge of branched amino acid,
              method of isolating the same and utilization thereof
              Patent: JP 2001169788-A 2 26-0UN-2001;
              DEBUSA HUELS AG, FORSCHUNGSZENTRUM JUELICH GMBH
COMMENT       OS Corynebacterium glutamicum ATCC14752
              PD JP 2001169788-A/2
              PD 26-JUN-2001
              PR 24-OCT-2000 JP 2000324315
              PR 27-OCT-1999 DE 19951708.8
              PI NICOLE KERNERUKUNEHITO, HERMANN SAHN, LOTHAR EGGERING, WALTER PI
FEATURES
source       1..753
              /organism="unidentified"
              /mol_type="genomic DNA"
              /db_xref="taxon:32644"
ORIGIN
Query Match      59.2%; Score 753; DB 6; Length 753;
Best Local Similarity 100.0%; Pred. No. 3,3e-196;
Matches 753; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 101 GTGCAAAAAGCAAGATTTATTCAGCCCTGAGAGTGTGGCCATCCAGGAGCCCTG 160
DB 1 GTGCAAAAAGCAAGATTTATTCAGCCCTGAGAGTGTGGCCATCCAGGAGCCCTG 60
QY 161 GAACAGATGATTAAGTTATCGGCTACGAATCGCGCAAGTCTAATAAACCCTCCCT 220
DB 61 GAACAGATGATTAAGTTATCGGCTACGAATCGCGCAAGTCTAATAAACCCTCCCT 120
QY 221 GCTGCAAGTTGGGCAATGACCCGATGATTTGCTGTCTTGGTTATTAATAC 280
DB 121 GCTGCAAGTTGGGCAATGACCCGATGATTTGCTGTCTTGGTTATTAATAC 180
QY 281 GGTACGAATGAGGAGGAGCCCACTGTTTCCGGCTGATTTCCGGGCTCCACGAA 340
DB 181 GGTACGAATGAGGAGGAGCCCACTGTTTCCGGCTGATTTCCGGGCTCCACGAA 240
QY 341 ATGCTGATCGCCCTGTTGTTGGGCGAGGAGCCCTGGCGCATCGCGTCAACCA 400
DB 241 ATGCTGATCGCCCTGTTGTTGGGCGAGGAGCCCTGGCGCATCGCGTCAACCA 300
QY 401 TTGCTGTAACCTCGGCAAGATTCATGCGTTTCATTCGCGTGCATGTGTCAAA 460
DB 301 TTGCTGTAACCTCGGCAAGATTCATGCGTTTCATTCGCGTGCATGTGTCAAA 360
QY 461 AACCCCATGCGCGTTTCTATTGCTTTGCGGCTTATGACGAAGCTTACGAGTCACT 520
DB 361 AACCCCATGCGCGTTTCTATTGCTTTGCGGCTTATGACGAAGCTTACGAGTCACT 420
QY 521 GGGGCAAGGCGGAGGAGTGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 580
DB 421 GGGGCAAGGCGGAGGAGTGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 480
QY 581 TCTACTGAGTATTCGCGGCTCTACCGAGTGGCGCATGCAAGTTGATTCCTTTGAA 640
DB 481 TCTACTGAGTATTCGCGGCTCTACCGAGTGGCGCATGCAAGTTGATTCCTTTGAA 540
QY 641 ATTAAAGGCTCGAGTTGCGCCCTTGTCTCTCTTTGTCACGCTGACTTGAATTCGCG 700
DB 541 ATTAAAGGCTCGAGTTGCGCCCTTGTCTCTCTTTGTCACGCTGACTTGAATTCGCG 600
QY 701 CGAAGCAAAAAGAGATCCCTTCTGCTGCTGCGAGGTTGAGCTTCAACATTCCTCT 760

```

```

DB 601 CGAAGCAAAAAGAGATCCCTTCTGCTGCTGCGAGGTTGAGCTTCAACATTCCTCT 660
QY 761 GTGTAAATTCAGAGTCAAGCCCTATTGTCGGGCTGTGATCTTGGTCTGTTAAC 820
DB 661 GTGTAAATTCAGAGTCAAGCCCTATTGTCGGGCTGTGATCTTGGTCTGTTAAC 720
QY 821 ATCCGATCTTTCTTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 853
DB 721 ATCCGATCTTTCTTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 753
RESULT 14
BD162490
LOCUS      BD162490              753 bp      DNA      linear      PAT 17-JAN-2003
DEFINITION Novel polynucleotide.
ACCESSION  BD162490
VERSION    BD162490.1 GI:27868248
KEYWORDS  JP 2002191370-A/289.
SOURCE     unidentified
ORGANISM   unidentified
            unclassified.
REFERENCE  1 (bases 1 to 753)
AUTHORS   Nakagawa, S., Mizoguchi, H., Ando, S., Hayashi, M., Ochiai, K.,
            Yokoi, H., Tateishi, N., Senoo, A., Ikeda, M. and Ozaki, A.
TITLE      Novel polynucleotide
JOURNAL    Patent: JP 2002191370-A 289 09-JUN-2002;
            KYOMA HAKKO KOGYO CO LTD
COMMENT    OS Corynebacterium glutamicum
            PN JP 2002191370-A/289
            PD 09-JUN-2002
            PF 15-DEC-2000 JP 2000405096
            PI SATOSHI NAKAGAWA, HIROSHI MIZOGUCHI, SEIKO ANDO, MIKIO HAYASHI,
            KEIKO OCHIAI,
            HARUHIKO YOKOI, NAOKO TATEISHI, AKIHINO SENOO, MASATO IKEDA, AKIO
            OZAKI
            PC C12N15/09, C12N15/09, C07K14/34, C07K16/12, C07K16/40, C12M1/00, PC
            C12N1/15,
            PC C12N1/19, C12N1/21, C12N5/10, C12N9/00, C12N9/02, C12P7/40, C12P13/
            PC 04, C12P13/08,
            PC C12P19/00, C12P19/34, C12P21/02, C12Q1/37, C12Q1/68, G01N33/53, PC
            G01N33/566
            PC G01N33/569, G01N33/68, G01N37/00, C12P21/08, C12N1/21, C12R1/15,
            PC C12N1/21, C12R1/13, C12N1/21, C12R1/01, C12P13/08, C12R1/15,
            PC C12N15/00,
            CC Novel polynucleotide
            CC Novel polynucleotide
            FH Key
            FT source
FEATURES
source       1..753
              /organism="unidentified"
              /mol_type="genomic DNA"
              /db_xref="taxon:32644"
ORIGIN
Query Match      59.2%; Score 753; DB 6; Length 753;
Best Local Similarity 100.0%; Pred. No. 3,3e-196;
Matches 753; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 101 GTGCAAAAAGCAAGATTTATTCAGCCCTGAGAGTGTGGCCATCCAGGAGCCCTG 160
DB 1 GTGCAAAAAGCAAGATTTATTCAGCCCTGAGAGTGTGGCCATCCAGGAGCCCTG 60
QY 161 GAACAGATGATTAAGTTATCGGCTACGAATCGCGCAAGTCTAATAAACCCTCCCT 220
DB 61 GAACAGATGATTAAGTTATCGGCTACGAATCGCGCAAGTCTAATAAACCCTCCCT 120
QY 221 GCTGCAAGTTGGGCAATGACCCGATGATTTGCTGTCTTGGTTATTAATAC 280
DB 121 GCTGCAAGTTGGGCAATGACCCGATGATTTGCTGTCTTGGTTATTAATAC 180
QY 281 GGTACGAATGAGGAGGAGCCCACTGTTTCCGGCTGATTTCCGGGCTCCACGAA 340

```


Db	663	GTCAATTACTTTGGCGCTCCGGGCGGTTCGGTCTTAATCCTTAAGCCCTCAAGTGATCA	604
QY	961	CAATTGTGGCAAAATGGCGATGGATGCGACGAGGAATCCTTGCCATTTTGACGCGA	1020
Db	603	CAATTGTGGGCAAAATGGCGATGGATGCGACGAGGAATCCTTGCCATTTTGACGCGA	544
QY	1021	TCAAGTTTGGCAGCAATGCGATGATCTGAAGACTTAACTTGGCTCATTTGCCGTT	1080
Db	543	TCAAGTTTGGCAGCAATGCGATGATCTGAAGACTTAACTTGGCTCATTTGCCGTT	484
QY	1081	GCGATTACAGTGTGGGCGCACTCTTCTTGGCGGTGACGCACTCTTGTGAAGCGTTGGCGCT	1140
Db	483	GCGATTACAGTGTGGGCGCACTCTTCTTGGCGGTGACGCACTCTTGTGAAGCGTTGGCGCT	424
QY	1141	GGCAACCATCGTTTTGTGTGAATCTTTCTTAAACTGCATTAATTAACAAAAT	1200
Db	423	GGCAACCATCGTTTTGTGTGAATCTTTCTTAAACTGCATTAATTAACAAAAT	364
QY	1201	CCGCATGCGCTCAATTGAAGGGGATGCGGATTTTAAAGAACTGAG-AAAAAGCTTAA	1259
Db	363	CCGCATGCGCTCAATTGAAGGGGATGCGGATTTTAAAGAACTGAGAAAAAGGCTTAA	304
QY	1260	GGAGACAGCGCT 1271	
Db	303	GGAGACAGCGCT 292	

LOCUS	AR391954	753 bp	DNA	linear	PAT 18-DEC-2003
DEFINITION	Sequence 2 from patent US 6613545.				
ACCESSION	AR391954				
VERSION	AR391954.1	GI:40115725			
KEYWORDS					
SOURCE	Unknown.				
ORGANISM	Unknown.				
REFERENCE	Unclassified.				
AUTHORS	1 (bases 1 to 753)				
TITLE	Kemernecht,N., Sahm,H., Eggeling,L. and Pfeifferle,W.				
JOURNAL	Nucleotide sequences coding for the export of branched chain amino acids, process for the isolation thereof and use thereof				
FEATURES	Patent: US 6613545-A 2 02-SEP-2003;				
source	Location/Qualifiers				
	1..753				
	/organism="unknown"				
ORIGIN	/mol_type="genomic DNA"				
Query Match	59.2%	Score 753;	DB 6;	Length 753;	
Best Local Similarity	100.0%;	Pred. No. 3.3e-196;			
Matches 753;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;	
QY	101	GTGCAAAAACCGAAGATTCATTCAACCCGAGGTCGTGGCATCCAGGACGCCCTG	160		
DB	1	GTGCAAAAACCGAAGATTCATTCAACCCGAGGTCGTGGCATCCAGGACGCCCTG	60		
QY	161	GACCAAGATGATTAAGATTATCGGCGCTACGAATCGCGCAAGGTCTAAAAAACCCTCCTT	220		
DB	61	GACCAAGATGATTAAGATTATCGGCGCTACGAATCGCGCAAGGTCTAAAAAACCCTCCTT	120		
QY	221	GCTGACGTTTGGGCAATGACCGAATTGATATTGGTCTCTTGGTTATTCAATAC	280		
DB	121	GCTGACGTTTGGGCAATGACCGAATTGATATTGGTCTCTTGGTTATTCAATAC	180		
QY	281	GGCTACGAATGTGTGGCAGCCGCCACTGTTTTCGGGCTGATTTTCGGGGCTTCACCGAA	340		
DB	181	GGCTACGAATGTGTGGCAGCCGCCACTGTTTTCGGGCTGATTTTCGGGGCTTCACCGAA	240		
QY	341	ATGCTGGTATCGCCCTCTGTTGGGGCGACGCCCTGGGGCGCATTCGGCTCACAA	400		
DB	341	ATGCTGGTATCGCCCTCTGTTGGGGCGACGCCCTGGGGCGCATTCGGCTCACAA	400		
QY	401	ATGCTGGTATCGCCCTCTGTTGGGGCGACGCCCTGGGGCGCATTCGGCTCACAA	460		
DB	401	ATGCTGGTATCGCCCTCTGTTGGGGCGACGCCCTGGGGCGCATTCGGCTCACAA	460		

QY	401	TTTGCTGTGAACCTTCGCGCACGTAATCTGAAGCGTTTTCATTCGCGCTGCATGTGGTCAAA	460
Db	301	TTTGCTGTGAACCTTCGCGCACGTAATCTGAAGCGTTTTCATTCGCGCTGCATGTGGTCAAA	360
QY	461	AACCCCATTTGCGCGTTTCTAATTCGGTTTTGGCGCTTATCGACGAAAGCTTAGCGAGTCACT	520
Db	361	AACCCCATTTGCGCGTTTCTAATTCGGTTTTGGCGCTTATCGACGAAAGCTTAGCGAGTCACT	420
QY	521	GCGGCGAGGCGCGGAGGCTGTGTGGCGGTGGCGACTTATCTCAATGCAAAATAGCGTTTCAC	580
Db	421	GCGGCGAGGCGCGGAGGCTGTGTGGCGGTGGCGACTTATCTCAATGCAAAATAGCGTTTCAC	480
QY	581	TCCTACTGGGTATTCGGCGGTCTTCACCGGAGTGGCGATTCGAGATTTCTCTTTGAA	640
Db	481	TCCTACTGGGTATTCGGCGGTCTTCACCGGAGTGGCGATTCGAGATTTCTCTTTGAA	540
QY	641	ATTAAAGGCGCTCGAGTTCGCGCTTGGCTGTCTCTTTTATCAAGCGACTTTGGATTCCTGCG	700
Db	541	ATTAAAGGCGCTCGAGTTCGCGCTTGGCTGTCTCTTTTATCAAGCGACTTTGGATTCCTGCG	600
QY	701	CGAAGCAAAAAGCAGATCCCTTCTCTGCTGTGCGACGTTTGAAGCTTCACCATTTGCTCTT	760
Db	601	CGAAGCAAAAAGCAGATCCCTTCTCTGCTGTGCGAGTTTGAAGCTTCACCATTTGCTCTT	660
QY	761	GTGGTAAATCCAGGTCAGGCCCTAATTTTGGCGCGCTGCGATCTCTTGGGCTGTGTGAC	820
Db	661	GTGGTAAATCCAGGTCAGGCCCTAATTTTGGCGCGCTGCGATCTCTTGGGCTGTGTGAC	720
QY	821	ATCCGTAATCTCTTGGGAAAGGCTGTAAA	853
Db	721	ATCCGTAATCTCTTGGGAAAGGCTGTAAA	753

LOCUS	AX120373	753 bp	DNA	linear	PAT 11-MAY-2001
DEFINITION	Sequence 289 from Patent EP1108790.				
ACCESSION	AX120373				
VERSION	AX120373.1	GI:14037088			
KEYWORDS					
SOURCE	Corynebacterium glutamicum				
ORGANISM	Corynebacterium glutamicum				
REFERENCE	Bacteria; Actinobacteriales; Corynebacteriaceae; Corynebacterium.				
AUTHORS	1 Nakagawa,S., Mizoguchi,H., Ando,S., Hayashi,M., Ochiai,K., Yokoi,H., Tateishi,N., Senoh,A., Ikeda,M. and Ozaki,A.				
TITLE	Novel polynucleotides				
JOURNAL	Patent: EP 1108790-A 289 20-JUN-2001;				
FEATURES	KYOMA HAKKO KOGYO CO., LTD. (JP)				
source	location/Qualifiers				
	1..753				
	/organism="Corynebacterium glutamicum"				
	/mol_type="unassigned DNA"				
	/db_xref="taxon:1718"				
ORIGIN					
Query Match	59.2%;	Score 753;	DB 6;	Length 753;	
Best Local Similarity	100.0%;	Pred. No. 3.3e-196;			
Matches 753;	Conservative	0;	Mismatches	0;	Indels 0; Gaps 0
QY	101 GTGCAAAAACGCAAGATTCATTCAGCTGAGGTGTGCCCATCCAAAGCAGCCCTG	160			
DB	1 GTGCAAAAACGCAAGATTCATTCAGCTGAGGTGTGCCCATCCAAAGCAGCCCTG	60			
QY	161 GAACAGATGATAAGCTTATCGGGCGATACGAAATCGGCGAAGGCTATAAAACCTCCCTT	220			
DB	61 GAACAGATGATAAGCTTATCGGGCGATACGAAATCGGCGAAGGCTATAAAACCTCCCTT	120			
QY	221 GCTGAGGTTGGGATGATGACCCGATTGGTATTTGGCTTCTTGGTTATTCAAATC	280			
DB	121 GCTGAGGTTGGGATGATGACCCGATTGGTATTTGGCTTCTTGGTTATTCAAATC	180			

QY 841 AAGGCTGCTAAATGACAACTGATTTCTCTGATTTCTCTGTTGTGCGAGATGTGCA 900
 Db 277569 AAGGCTGCTAAATGACAACTGATTTCTCTGATTTCTCTGTTGTGCGAGATGTGCA 277628
 QY 901 GTGATTAATTTTCCGCTCCGGGGGGTTCGGTCTTAATCTTAACCCCTACGTGATCA 960
 Db 277629 GTGATTAATTTTCCGCTCCGGGGGGTTCGGTCTTAATCTTAACCCCTACGTGATCA 277688
 QY 961 CAATTGTGGGCAAAATGCGATGTGATGCGACAGAGATCTTGCAATTTTGACCGCA 1020
 Db 277689 CAATTGTGGGCAAAATGCGATGTGATGCGACAGAGATCTTGCAATTTTGACCGCA 277748
 QY 1021 TCAAGCTTTGCGAGCAATGAGATCTGAAAGACTTAACTTTGGTCTCAATGGCGCT 1080
 Db 277749 TCAAGCTTTGCGAGCAATGAGATCTGAAAGACTTAACTTTGGTCTCAATGGCGCT 277808
 QY 1081 GCGATTACAGTGTGGCGCATCTTCTTGGCGGTGCGACGACCTTGTGAGCGTGGCGCT 1140
 Db 277809 GCGATTACAGTGTGGCGCATCTTCTTGGCGGTGCGACGACCTTGTGAGCGTGGCGCT 277868
 QY 1141 GGCACCATGCTTTTGTGGAATGCTGATCTTTCTTAACTGATTAATTAACAAAT 1200
 Db 277869 GGCACCATGCTTTTGTGGAATGCTGATCTTTCTTAACTGATTAATTAACAAAT 277928
 QY 1201 CCGCATGCTCTCAATTGGAAGGGATGCGAATTTTAAAGAACTTAAAGGCTTAAG 1260
 Db 277929 CCGCATGCTCTCAATTGGAAGGGATGCGAATTTTAAAGAACTTAAAGGCTTAAG 277988
 QY 1261 CAGACAGCGCT 1271
 Db 277989 CAGACAGCGCT 277999

RESULT 8
 AX120085 349980 bp DNA linear PAT 11-MAY-2001
 LOCUS AX120085
 DEFINITION Sequence 1 from Patent EP1108790.
 ACCESSION AX120085 AX114121
 VERSION AX120085.1 GI:14036800
 KEYWORDS
 SOURCE Corynebacterium glutamicum
 ORGANISM Corynebacterium glutamicum
 Bacteria; Actinobacteria; Actinobacteriales; Actinomycetales;
 Corynebacteriaceae; Corynebacterinae; Corynebacterium.
 REFERENCE 1 Nakagawa, S., Mizoguchi, H., Ando, S., Hayashi, M., Ochiai, K.,
 Yoko, H., Tateishi, N., Senoh, A., Ikeda, M. and Ozaki, A.,
 Novel polynucleotides
 Patent: EP 1108790-A 1 20-JUN-2001,
 KYOMA HAKKO KOGYO CO., LTD. (JP)
 FEATURES
 source location/Qualifiers
 1..349980
 /organism="Corynebacterium glutamicum"
 /mol_type="genomic DNA"
 /db_xref="taxon:1718"
 /note="Seq 1 to long (3,309,400) split in 11, seq 1
 0,000,001 0,349,980 349,980"

ORIGIN
 Query Match 100.0%; Score 1271; DB 6; Length 349980;
 Best Local Similarity 100.0%; Fred. No. 0;
 Matches 1271; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 276849 TCATTCAAGCTGAGAGTGTCCCATCCAGAGCAGCCCTGGAACCAATGATTAAGTTA 276908
 QY 181 TCGGCGCTACGAATCGCGCAGAGCTTAAACCTCCCTGCTGAGGTTGGGATGTA 240
 Db 276909 TCGGCGCTACGAATCGCGCAGAGCTTAAACCTCCCTGCTGAGGTTGGGATGTA 276968
 QY 241 CCGGATTTGATTTGGCTTTGCTCTTGGTATTCAATACGCTACGATGATGATGAGCAGC 300
 Db 276969 CCGGATTTGATTTGGCTTTGCTCTTGGTATTCAATACGCTACGATGATGATGAGCAGC 277028
 QY 301 CCGGATTTGATTTGGCTTTGCTCTTGGTATTCAATACGCTACGATGATGATGAGCAGC 360
 Db 277029 CCGGATTTGATTTGGCTTTGCTCTTGGTATTCAATACGCTACGATGATGATGAGCAGC 277088
 QY 361 TGTGGCGGAGCGCCCTGCGGCGCATGCGCTACCAATTTGCTGTAACCTTCGCA 420
 Db 277089 TGTGGCGGAGCGCCCTGCGGCGCATGCGCTACCAATTTGCTGTAACCTTCGCA 277148
 QY 421 CGGATTTGATTTGGCTTTGCTCTTGGTATTCAATACGCTACGATGATGATGAGCAGC 480
 Db 277149 CGGATTTGATTTGGCTTTGCTCTTGGTATTCAATACGCTACGATGATGATGAGCAGC 277208
 QY 481 TTGGGTTTTCGGCTTATGCAAGAACTTACGAGCTCACTGCGCGCAGCGCGAGCTG 540
 Db 277209 TTGGGTTTTCGGCTTATGCAAGAACTTACGAGCTCACTGCGCGCAGCGCGAGCTG 277268
 QY 541 GTGGGCTGCGCATTTATGCAATGCAATACGCTTACCTTCACTGCGGATTTCCGCGG 600
 Db 277269 GTGGGCTGCGCATTTATGCAATGCAATACGCTTACCTTCACTGCGGATTTCCGCGG 277328
 QY 601 TCTACCGGAGTGGCGATGCGAGAGTTATCTTTTGAATTAAGGCGCTCGAGTGC 660
 Db 277329 TCTACCGGAGTGGCGATGCGAGAGTTATCTTTTGAATTAAGGCGCTCGAGTGC 277388
 QY 661 CTTTGTCTCTCTTTTGTGACGCTGACTTGTGATCTTCTGCGAGCAAGAAAGCGATGCC 720
 Db 277389 CTTTGTCTCTCTTTTGTGACGCTGACTTGTGATCTTCTGCGAGCAAGAAAGCGATGCC 277448
 QY 721 TTCTGTCTCTCTGCGAGTTTGTGAGCTTCAACATGCTCTTGTGATTAATTCAGGTCAGC 780
 Db 277449 TTCTGTCTCTCTGCGAGTTTGTGAGCTTCAACATGCTCTTGTGATTAATTCAGGTCAGC 277508
 QY 781 CCGATTTGGCGGCGCTGATCTTCTGCGTGTGTAACGCGGATCTTCTTGGG 840
 Db 277509 CCGATTTGGCGGCGCTGATCTTCTGCGTGTGTAACGCGGATCTTCTTGGG 277568
 QY 841 AAGGCTGCTAAATGACAACTGATTTCTCTGATTTCTCTGTTGTGCGAGATGTGCA 900
 Db 277569 AAGGCTGCTAAATGACAACTGATTTCTCTGATTTCTCTGTTGTGCGAGATGTGCA 277628
 QY 901 GTGATTAATTTTCCGCTCCGGGGGGTTCGGTCTTAATCTTAACCCCTACGTGATCA 960
 Db 277629 GTGATTAATTTTCCGCTCCGGGGGGTTCGGTCTTAATCTTAACCCCTACGTGATCA 277688
 QY 961 CAATTGTGGGCAAAATGCGATGTGATGCGACAGAGATCTTGCAATTTTGACCGCA 1020
 Db 277689 CAATTGTGGGCAAAATGCGATGTGATGCGACAGAGATCTTGCAATTTTGACCGCA 277748
 QY 1021 TCAAGCTTTGCGAGCAATGAGATCTGAAAGACTTAACTTTGGTCTCAATGGCGCT 1080
 Db 277749 TCAAGCTTTGCGAGCAATGAGATCTGAAAGACTTAACTTTGGTCTCAATGGCGCT 277808
 QY 1081 GCGATTACAGTGTGGCGCATCTTCTTGGCGGTGCGACGACCTTGTGAGCGTGGCGCT 1140
 Db 277809 GCGATTACAGTGTGGCGCATCTTCTTGGCGGTGCGACGACCTTGTGAGCGTGGCGCT 277868
 QY 1141 GGCACCATGCTTTTGTGGAATGCTGATCTTTCTTAACTGATTAATTAACAAAT 1200
 Db 277869 GGCACCATGCTTTTGTGGAATGCTGATCTTTCTTAACTGATTAATTAACAAAT 277928
 QY 1201 CCGCATGCTCTCAATTGGAAGGGATGCGAATTTTAAAGAACTTAAAGGCTTAAG 1260
 Db 277929 CCGCATGCTCTCAATTGGAAGGGATGCGAATTTTAAAGAACTTAAAGGCTTAAG 277988


```

RKKRAGSLDGRQKRYVGAESLSVANKELQKRGKQDIAGCVTSNMEELVAKIA
OHTREVMKDKKLPITCDSTAMATNLRMQRQIQLVIAEKVPIITELMIFGQAPS
WRKGLHFKRGSPRTYG"
/gene="Cg10006"
5354. 7489
5354. /gene="Cg10006"
5354. 7489
/gene="Cg10006"
/EC_number="5.99.1.3"
/Note="PF00204:DNA topoisomerase II (N-terminal region)
PF00986:DNA gyrase B subunit, carboxyl terminus
PF01751:Toprim domain
PF02518:Histidine kinase-, DNA gyrase B-, phytochrome-like
ATPase
TIGR01055:pair_Gneg: DNA topoisomerase IV, B subunit
TIGR01058:pair_Gpos: DNA topoisomerase IV, B subunit
TIGR01059:gyrB: DNA gyrase, B subunit"
/codon_start=1
/translation="DNA gyrase (topoisomerase II) B subunit"
/protein_id="BAB97399.1"
/db_xref="GI:21322770"
/translation="NRGTTMGPKRYRMRKLYRISSEESLVANTHENYDASITILE
GLEAVRKPGMTIGSTGPRGLHLIWEVVDNSVDEAMAGHATKVEVTLLEPGQVTD
DORGI.PVMEHPSGAPVQVMTQLHAGKFDSDSYAGSLHVGISVNLSTREVA
DILKHKWYQNFERSVDELEEGNAGGTGTRTFMDAEI FETTERDEPTISRLO
EAAFLNKGITLTLDNRATDELELEALAEQGETATSELEIDNETELVEETDAPK
KPKRKKKI.PHYNGLEDVYHVRNSKTNPISVSEPAQGDHDEVAQAOMSSYK
ESVHPFANTIRREGTEGERSALSTLMRPAEHLLEKEMNLTDGDCRGLSA
VTSVAGDPQEGQRTKIKNTKIKS.FIQRANNEIGMILANPBAKVIINRKYGA
QNRLLARKARDLVRRKSAVDLGLPGLADCKSKDPESELTIVEGDAGSASAGSD
SMFOALPLRGIKLVNERKRLDKLVKAEVQALITALTGTHDEIDINKLHYKILM
AADVDGQHIATLTLTLFRFMPDLVABGHVLAQPLPKWKMGEGEPFASBERD
EOLNEGLAGRKINDQGIQRYKGLGEMNASLMEETMDPTVRLIRVDITDAQRADE
LFSIMGDVVARSPITRNADVRFDI"
7830. 8798
/gene="Cg10007"
7830. 8798
/gene="Cg10007"
/codon_start=1
/translation="Hypothetical protein"
/product="Hypothetical protein"
/protein_id="BAB97400.1"
/db_xref="GI:21322771"
/translation="KIQITILKIQSIKAKQRMATTLITDILGHVSTETPTAHV
LPDGSSTPVQIWDNDKDSQVWLMWPGFMGYRPLAALNKGPHVAGELRGQG
OSAKAKRSQWGYHDLASVDFPQIAAKKALDEBEGHMFSLHSVGGQISCLFAA
REBAKYNLRALFVGAGSPFRPTSPKRGKLGAVLLGIGIIGHIVGFWPKVLGK
DIIVGROSGTHMRWRFRHKNSLDLTAODINVEVWKVSIPIITSRCDDDDCP
QASIDGLAFVPAQIKMIEIPDALGHNRMAREPSPYVLFLEQAL"
complement (8795. 9466)
/gene="Cg10008"
complement (8795. 9466)
/gene="Cg10008"
/codon_start=1
/translation="Hypothetical membrane protein"
/product="Hypothetical membrane protein"
/protein_id="BAB97401.1"
/db_xref="GI:21322772"
/translation="MSNENKRLPDPAQRSLPTRNAMPVVGIFIMSFAIGTVHAGLG
TTISLPLVMTAGSLSGWTTIFNGSMILCOITLVLSKPTQALQVILALFGLF
LDDLSLQLTWAGTDNYPVAMIMVIVSTILMSIGIIVLPNITPILAGEIVSALVK
FVVERGTMKQIVDMTFVSVAALISWITMGILIGVREGTFAAFIFGFEVROWRLYL
RSIGH"
complement (9471. 9914)
/gene="Cg10009"
complement (9471. 9914)
/gene="Cg10009"
/codon_start=1
/translation="Hypothetical protein"
/product="Hypothetical protein"
/protein_id="BAB97402.1"
/db_xref="GI:21322773"
/translation="MTSLRDLVSTNADAAIADMSAFIEETVKGOSGIGIALKQAMG

```

```

AATKIDSDIVAKSRRLPETADSLDGLMOEYONGTAPADFGALFANSSTALDAIS
VADKNAEBSINVRGLKRYKGVGRKAAKYLBOELPRLIGLIEBNAQ"
complement (10104. 11177)
/gene="Cg10010"
complement (10104. 11177)
/gene="Cg10010"
/codon_start=1
/translation="Hypothetical protein"
/product="Hypothetical protein"
/protein_id="BAB97403.1"
/db_xref="GI:21322774"
/translation="MEMLRQRLPIHTAPHLVPAVRVPHQIRATDADRTSLQRDLLELS

Query Match      100.0% Score 1271; DB 1; Length 34000;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1271; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 GCGCGATCAATGGAATCTAGCTTCATATATTGACAAATAGCTTAGTGAAGTCCGAAAC 60
Db      276729 GCGCGATCAATGGAATCTAGCTTCATATATTGACAAATAGCTTAGTGAAGTCCGAAAC 276788
QY      61 TGGCAACAAACTACCGGCAATGTGTGATGATGTAGTGTGCAAAAACGCAAGAT 120
Db      276789 TGGCAACAAACTACCGGCAATGTGTGATGATGTAGTGTGCAAAAACGCAAGAT 276848
QY      121 TCAATTCAAGCCTGAGAGTGTGGCCATTCAGAGCCTCTGAAACGATGATTAAGGTTA 180
Db      276849 TCAATTCAAGCCTGAGAGTGTGGCCATTCAGAGCCTCTGAAACGATGATTAAGGTTA 276908
QY      181 TCGCGCTAGAAATGCGCGCAAGGTCTAAACCTCCCTGCTGCGAGGTTTGGGCATGTA 240
Db      276909 TCGCGCTAGAAATGCGCGCAAGGTCTAAACCTCCCTGCTGCGAGGTTTGGGCATGTA 276968
QY      241 CCGGATGTATTCGTTTGGTCTCTGTGTTATTCATAACGCTACGAATGTGGGCAAGC 300
Db      276969 CCGGATGTATTCGTTTGGTCTCTGTGTTATTCATAACGCTACGAATGTGGGCAAGC 277028
QY      301 CCCACTGTTTTCGGGCTGATTTTTCGGGGCTCCACGAAATGCTGTGATCGCCCTCGT 360
Db      277029 CCCACTGTTTTCGGGCTGATTTTTCGGGGCTCCACGAAATGCTGTGATCGCCCTCGT 277088
QY      361 TGTGGGCGCAGCGCCCTGCGGCGCAATCGCGCTCACCAATTGCTGTGTAATCTCGCCA 420
Db      277089 TGTGGGCGCAGCGCCCTGCGGCGCAATCGCGCTCACCAATTGCTGTGTAATCTCGCCA 277148
QY      421 CGTATTCTATGCGTTTTCATTTCCCGCTGCATGTGTGTCMAAAACCCATTGCCGTTCTA 480
Db      277149 CGTATTCTATGCGTTTTCATTTCCCGCTGCATGTGTGTCMAAAACCCATTGCCGTTCTA 277208
QY      481 TTGGTTTTGCGGCTTATCGACGAGCGCTACGAGTCACTGCGGCGCAGCCGAGGCTG 540
Db      277209 TTGGTTTTGCGGCTTATCGACGAGCGCTACGAGTCACTGCGGCGCAGCCGAGGCTG 277268
QY      541 GTCGGCGTGGGACTTATCTCAATGCAATAGCGTTTCACTCTACTGCGTATTTCCGCG 600
Db      277269 GTCGGCGTGGGACTTATCTCAATGCAATAGCGTTTCACTCTACTGCGTATTTCCGCG 277328
QY      601 TCTCACCAGGAGTGGCAATCGCAGAGTGAATCTCTTGAATTAAGGCGCTCGAGTTGCG 660
Db      277329 TCTCACCAGGAGTGGCAATCGCAGAGTGAATCTCTTGAATTAAGGCGCTCGAGTTGCG 277388
QY      661 CTTTGGCTCTCTCTTTGTACAGCTGACTTGTGATTTCTGCGCGAAGCAAAAGCATGCC 720
Db      277389 CTTTGGCTCTCTCTTTGTACAGCTGACTTGTGATTTCTGCGCGAAGCAAAAGCATGCC 277448
QY      721 TTTCTGTGCTGTCGACAGTTTGAAGCTCACCATTTGCTCTTGTGTAATTCAGGTCAGGC 780
Db      277449 TTTCTGTGCTGTCGACAGTTTGAAGCTCACCATTTGCTCTTGTGTAATTCAGGTCAGGC 277508
QY      781 CCTATTGCGGCGGTGTGATCTTCTTGGGCTGTGTGACCATCCGGTACTTCTTCTGGG 840
Db      277509 CCTATTGCGGCGGTGTGATCTTCTTGGGCTGTGTGACCATCCGGTACTTCTTCTGGG 277568

```

QY 901 GTCATTACTTTGGCGCTCCGGGCGGTCCGTTCTTAATCCTTAAGCCCTACGTAATCA 960
 DB 901 GTCATTACTTTGGCGCTCCGGGCGGTCCGTTCTTAATCCTTAAGCCCTACGTAATCA 960
 QY 961 CAATTGTGGGCAAAATGGCGATGTGATGCGACAGGAATCCTTGCCATTTTACCGCA 1020
 DB 961 CAATTGTGGGCAAAATGGCGATGTGATGCGACAGGAATCCTTGCCATTTTACCGCA 1020
 QY 1021 TCAAGTTTCCGAGCAATATGCAATATCTGAAGACTTAACTTTGCTCTGATGCGGCT 1080
 DB 1021 TCAAGTTTCCGAGCAATATGCAATATCTGAAGACTTAACTTTGCTCTGATGCGGCT 1080
 QY 1081 GCGATTACAGTGTGGCGCATCTTCTTGCGCGTGCACGACCTTGTTAGCGGCT 1140
 DB 1081 GCGATTACAGTGTGGCGCATCTTCTTGCGCGTGCACGACCTTGTTAGCGGCT 1140
 QY 1141 GCGACCAATGTTTGTGTTGAGTGTGATCTTTTCTAAAAGTGCATTAATTAACAAAAT 1200
 DB 1141 GCGACCAATGTTTGTGTTGAGTGTGATCTTTTCTAAAAGTGCATTAATTAACAAAAT 1200
 QY 1201 CCGCATGCGCCCAATTTGAAGGGGATGCGGATTTTGAAGAACTAGAAAAGGCTTAAG 1260
 DB 1201 CCGCATGCGCCCAATTTGAAGGGGATGCGGATTTTGAAGAACTAGAAAAGGCTTAAG 1260
 QY 1261 CAGACAGCGCT 1271
 DB 1261 CAGACAGCGCT 1271

RESULT 7
 LOCUS AP005274 340000 bp DNA linear ECT 08-AUG-2002
 DEFINITION Corynebacterium glutamicum ATCC 13032 DNA, complete genome, section 1/10.
 ACCESSION AP005274 BA000036
 VERSION AP005274.1 GI:21322764
 KEYWORDS
 ORGANISM
 Corynebacterium glutamicum ATCC 13032
 Corynebacterium glutamicum ATCC 13032
 Bacteria; Actinobacteria; Actinobacteriales; Corynebacteriaceae; Corynebacterium.
 REFERENCE
 1 Nakagawa, S.
 Complete genomic sequence of Corynebacterium glutamicum ATCC 13032
 TITLE Unpublished
 JOURNAL 2 (bases 1 to 340000)
 AUTHORS Nakagawa, S.
 DIRECT SUBMISSION
 JOURNAL Submitted (24-MAY-2002) Satoshi Nakagawa, Kyowa Hakko Kogyo Co.
 TITLE Bacteria; Actinobacteria; Actinobacteriales; Corynebacteriaceae; Corynebacterium.
 JOURNAL Tokyo 194-8533, Japan (E-mail:snakagawa@kaken.gen.com, Tel:81-44-829-3031, Fax:81-44-813-1651)
 COMMENT This sequence is conducted by collaboration of Kyowa Hakko Kogyo Co. Ltd. And Kitasato University.
 FEATURES
 source location/Qualifiers
 1. 340000
 /organism="Corynebacterium glutamicum ATCC 13032"
 /mol_type="genomic DNA"
 /strain="ATCC 13032"
 /db_xref="taxon:196627"
 /note="ATCC 13032"
 1. 1575
 /gene="Cg10001"
 1. 1575
 /gene="Cg10001"
 /note="PB00308: Bacterial dna protein
 TIGR00362: DnaA: chromosomal replication initiator protein
 DnaA"
 /codon_start=1
 /transl_table=11
 /product="Hypothetical protein"
 /protein_id="BAB97394.1"
 /db_xref="GI:21322765"

gene
 CDS
 gene
 CDS

/translation="MSONSSSLEETWROVADLTLLSQDADSGPDLPTPTORAYLNT
 KPIAIVDGYAVSTPNAMAKNYIENDLGLALFRVLSLRGSPSLAVSEPOEIPET
 PAOEFKXPDPAPVSISSNKAPOKOVGREGASTDGMERTSAPAPERPAPIDPE
 PELATFORIPRETTPAHNRREVSILNPKTFSEFVIGPPRFANAAVAASAPAFN
 PLTISGSGSLGTHLLHAGVNTAQELQPELRITKYSSEFTNDYINSRVDQCEFR
 RYRNLDILMWDDIOFLAGEKEQEEFFFNALHQADKOIILSDRPPKQITLEDRL
 RTRPEGLITDIQPDLETRIALIMKKQIDGTHTVREVELIASRFESSIRLEGLA
 IRVSAESSLINOPIDKEMALVALRDILPEPEMEITAPYIMEVTAIEFEISVDTLGA
 GTRAVAHAROLAMYLCRELTDMSLPKIGDVFGGHDHTVWVADKRIQEMTEKDY
 DELQUTOLIKERGN"
 complement (1594, 1920)
 /gene="Cg10002"
 /complement (1594, 1920)
 /gene="Cg10002"
 /codon_start=1
 /transl_table=11
 /product="Hypothetical protein"
 /protein_id="BAB97395.1"
 /db_xref="GI:21322766"
 /translation="MIVDFVYLWIDFKYRILSPVCADVSPAYPQFLSCAQGKFRG
 ITPVTKLTQLQCVNTAVNKKVDNFWAFTGSPQNDCTPKLQKQPSDEGSRRA
 ENLFFV"
 2292, 3476
 /gene="Cg10003"
 2292, 3476
 /gene="Cg10003"
 /EC_number="2.7.7.7"
 /note="PF00712: DNA polymerase III beta subunit, N-terminal
 domain
 PF02767: DNA polymerase III beta subunit, C-terminal domain
 PF02768: DNA polymerase III beta subunit, C-terminal domain
 TIGR00663: dnan: DNA polymerase III, beta subunit"
 /codon_start=1
 /transl_table=11
 /product="DNA polymerase III beta subunit"
 /protein_id="BAB97396.1"
 /db_xref="GI:21322767"
 /translation="MESQNVSPRVAREDLVTAVAWARNLPTRKQPVLRAMLITAD
 EGLEAGDYDSTRTVRLSABVSGPRTIVAGKLSLETGSLPNKRVDPRIQSKAFV
 TCSSVSPFELPLLDYPMVLPKIPAGTGSINAKLTFEAVSQVAPAKXDSLPMLTG
 SMEIVNQINLALTDRELRALRTFEWANDPELNVLLIPARTLDNASSLDSGLND
 IDIAGVTGQVGRGELFGVHTDRETTLRLLADPENILPLPKERTALASVEIALV
 DAIRVSLVARNAGIVLHFSBGQVLTAGATEAGAAERTLPQAFYKXELTIAFNQY
 LKQGLSVPTSRVAFGTSPRAIMIPPEMPSANENGIFQTPPTVFTYLLMPVRL
 PG"
 3585, 4769
 /gene="Cg10004"
 3585, 4769
 /gene="Cg10004"
 /note="TIGR00611: recf: RecF protein"
 /codon_start=1
 /transl_table=11
 /product="Recombination DNA repair ATPase"
 /protein_id="BAB97397.1"
 /db_xref="GI:21322768"
 /translation="WHIRSLRLDYRWPELKYDLBPGLTFVPIGRNGFGKTNIVEIG
 YLAHSSHRVSDAPLVRHAENARVSAVAVNGRELAHLILKPAANAQASINRKY
 RTRELIGVYKTLFAPEDIALYKGPASRRYLLDDIIRACPRMAGVADYKVKQ
 RNALIKTATILALRGVTEGGAALSTLDWDQGLRAGVVARAFALNEGPKTY
 EATTTAPBSRPAAVNYKTTIDQGLSOFSEPDGIIENITLTLAKROEITLIDVSE
 VGRHDDVDMLGSDQPAKGFASGRTSFRSLRIAPFNILKSDGDPDITLIDVSE
 LDAGRRKLVGIAQVEVQVLTAAVDDLPENILKVLTAQHTVTDPTGTGRLSLDV
 QP"
 4766, 5302
 /gene="Cg10005"
 4766, 5302
 /gene="Cg10005"
 /codon_start=1
 /transl_table=11
 /product="Hypothetical protein"
 /protein_id="BAB97398.1"
 /db_xref="GI:21322769"
 /translation="MTDPIEQAFERIABAMRANGVPLDKNDAPFRPPAPKGVEX

QY 721 TTCCTGCTGCTGCGAGGTTTGAAGCTTACCATGCTCTTGTGTAAATTCAGGTACAGGC 780
 DB 721 TTCCTGCTGCTGCGAGGTTTGAAGCTTACCATGCTCTTGTGTAAATTCAGGTACAGGC 780
 QY 781 CCTATTTCGGGCGCTGCTGATCTTCTTGGGCTGTGTGACCATCCGGTACTTCTTGGG 840
 DB 781 CCTATTTCGGGCGCTGCTGATCTTCTTGGGCTGTGTGACCATCCGGTACTTCTTGGG 840
 QY 841 AAAGGCTGCTAAATGACACTGATTTCTCTGTAATTCCTTGTGTGACAGTATGTGCA 900
 DB 841 AAAGGCTGCTAAATGACACTGATTTCTCTGTAATTCCTTGTGTGACAGTATGTGCA 900
 QY 901 GTGATTACTTTTGGCTCCGCGCGGCTCCGCTTCTTAATCTTAAAGCCCTTACGTGAATCA 960
 DB 901 GTGATTACTTTTGGCTCCGCGCGGCTCCGCTTCTTAATCTTAAAGCCCTTACGTGAATCA 960
 QY 961 CAATTGTGGGCAAAATGGCGCATGTGATGCGCAGAGGATCCTTGCCATTGTGACGCA 1020
 DB 961 CAATTGTGGGCAAAATGGCGCATGTGATGCGCAGAGGATCCTTGCCATTGTGACGCA 1020
 QY 1021 TCAAGCTTTCGAGCAATGCGATGATCTGAAGACTTAACTTGTGCTATTGCGCTT 1080
 DB 1021 TCAAGCTTTCGAGCAATGCGATGATCTGAAGACTTAACTTGTGCTATTGCGCTT 1080
 QY 1081 GCGATTACAGTGTGGCGCATCTTCTGCGGTGCAAGCACTTTTGAAGCGTGGCGCT 1140
 DB 1081 GCGATTACAGTGTGGCGCATCTTCTGCGGTGCAAGCACTTTTGAAGCGTGGCGCT 1140
 QY 1141 GGCACCATGCTTTTGTGTGACCTGTGATCTTTTCTAAACTGTCAATAATACAAAT 1200
 DB 1141 GGCACCATGCTTTTGTGTGACCTGTGATCTTTTCTAAACTGTCAATAATACAAAT 1200
 QY 1201 CCGCATGCTCTCAATTGAAAGGGAGATGCGGATTTTAAAGAACTTAAGAAAGGCTTAA 1260
 DB 1201 CCGCATGCTCTCAATTGAAAGGGAGATGCGGATTTTAAAGAACTTAAGAAAGGCTTAA 1260
 QY 1261 CAGACAGCGCT 1271
 DB 1261 CAGACAGCGCT 1271

RESULT 6
 BD014994 1271 bp DNA linear PAT 27-AUG-2002
 LOCUS Nucleotide sequence encoding the discharge of branched amino acid,
 DERIVATION method of isolating the same and utilization thereof.
 ACCESSION BD014994
 VERSION BD014994.1 GI:22555801
 KEYWORDS JP 200116978-A/4.
 SOURCE unidentified
 ORGANISM unclassified.
 1 (bases 1 to 1271)
 Kernerukunehito, N., Sahm, H., Eggering, L. and Pfeifferle, W.
 Nucleotide sequence encoding the discharge of branched amino acid,
 method of isolating the same and utilization thereof
 Patent: JP 200116978-A 4 26-JUN-2001;
 DEGUSA HUELS AG, FORSCHUNGSZENTRUM JUELICH GMBH
 OS Corynebacterium glutamicum ATCC13032
 PN JP 200116978-A/4
 PD 26-JUN-2001
 PR 24-OCT-2000 JP 2000324315
 PR 27-OCT-1999 DE 19951708.8
 PI NICOLE KERNERUKUNEHITO, HERMANN SAHM, LOTAR EGGERING, WALTER PI
 PFEIFFERLE
 PC C12N15/09, C07K14/34, C12N1/21, C12P13/06, C12P13/08, C12P13/08, C12N15/09, C12R1/15, (C12N1/21, C12R1/15), (C12P13/06, C12R1/15),
 PC C12N15/00, C12R1/15)
 PC C12N15/00, C12R1/15)
 CC bnf
 CC bnf
 FH key Location/Qualifiers

FEATURES
 source
 1..1271
 /organism="unclassified"
 /mol_type="genomic DNA"
 /db_xref="taxon:32644"

ORIGIN
 Query Match 100.0%; Score 1271; DB 6; Length 1271;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1271; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCGGATCAATGAAATCTAGCTTCATATATTTGACAAATAGCTAGTGGTGGCAAC 60
 DB 1 GCGGATCAATGAAATCTAGCTTCATATATTTGACAAATAGCTAGTGGTGGCAAC 60
 QY 61 TGGCAACAAATACCCGGCAATTGTGTATGATTTGTGCAAAAACCGAAGAGAT 120
 DB 61 TGGCAACAAATACCCGGCAATTGTGTATGATTTGTGCAAAAACCGAAGAGAT 120
 QY 121 TCAATCAAGCTGAGAGTGTGGCATCCAAAGGAGCCCTGGAACGATGATAAGGTTA 180
 DB 121 TCAATCAAGCTGAGAGTGTGGCATCCAAAGGAGCCCTGGAACGATGATAAGGTTA 180
 QY 181 TCGGCGCTACGAATGCGCGCAAGTCTAAAACCTCCCTTGCTGCAAGGTTGGGCAATGTA 240
 DB 181 TCGGCGCTACGAATGCGCGCAAGTCTAAAACCTCCCTTGCTGCAAGGTTGGGCAATGTA 240
 QY 241 CCGGATGATTTGCTGCTTGTGTGCTCTTGTGTTAATACGGCTAGAAATGTTGGGCAAG 300
 DB 241 CCGGATGATTTGCTGCTTGTGTGCTCTTGTGTTAATACGGCTAGAAATGTTGGGCAAG 300
 QY 301 CCGATTTGATTTGCTGCTTGTGTGCTCTTGTGTTAATACGGCTAGAAATGTTGGGCAAG 360
 DB 301 CCGATTTGATTTGCTGCTTGTGTGCTCTTGTGTTAATACGGCTAGAAATGTTGGGCAAG 360
 QY 361 TGTGGGCGAGCGCCCTGCGGCGCATGCGGCTCACCAATGCTGTGATCCGCGCA 420
 DB 361 TGTGGGCGAGCGCCCTGCGGCGCATGCGGCTCACCAATGCTGTGATCCGCGCA 420
 QY 421 CGATTTCTAGCGCTTTTCAATTCGCGCTGATGTGTCAAAAACCGATTGCCGTTCTA 480
 DB 421 CGATTTCTAGCGCTTTTCAATTCGCGCTGATGTGTCAAAAACCGATTGCCGTTCTA 480
 QY 481 TTGCGTTTTCGCTTATGACAGAGCTTACGAGTCACTGCGGCGCAAGCCGCAAGCTG 540
 DB 481 TTGCGTTTTCGCTTATGACAGAGCTTACGAGTCACTGCGGCGCAAGCCGCAAGCTG 540
 QY 541 GTGCGGCTGGGCACTTATCTCAATGCAATAGCGTTTCACTCTACTAGGTTATTCGCGG 600
 DB 541 GTGCGGCTGGGCACTTATCTCAATGCAATAGCGTTTCACTCTACTAGGTTATTCGCGG 600
 QY 601 TCTCAGCGAGTGGCGATGCGAGAGTGTATCTTTTGAATTAAGGCTCGAGTTGCG 660
 DB 601 TCTCAGCGAGTGGCGATGCGAGAGTGTATCTTTTGAATTAAGGCTCGAGTTGCG 660
 QY 661 CTTTGTGCTCTCTTGTACAGCGTGAAGTTGATTCCTGCGGAAAGAAAGAGATCCC 720
 DB 661 CTTTGTGCTCTCTTGTACAGCGTGAAGTTGATTCCTGCGGAAAGAAAGAGATCCC 720
 QY 721 TTCCTGCTGCTGCGAGGTTTGAAGCTTACCATGCTCTTGTGTAAATTCAGGTACAGGC 780
 DB 721 TTCCTGCTGCTGCGAGGTTTGAAGCTTACCATGCTCTTGTGTAAATTCAGGTACAGGC 780
 QY 781 CCTATTTCGGGCGCTGCTGATCTTCTTGGGCTGTGTGACCATCCGGTACTTCTTGGG 840
 DB 781 CCTATTTCGGGCGCTGCTGATCTTCTTGGGCTGTGTGACCATCCGGTACTTCTTGGG 840
 QY 841 AAAGGCTGCTAAATGACACTGATTTCTCTGTAATTCCTTGTGTGACAGTATGTGCA 900
 DB 841 AAAGGCTGCTAAATGACACTGATTTCTCTGTAATTCCTTGTGTGACAGTATGTGCA 900

Db 481 TTGGTTTCGCGCTTATCGACGAGCGTATCGACGCGCCGCGAGGCTG 540
 QY 541 GTGGCGGTGGGCACTTATCTCAATGCAATAGCGTTCACTCTCACTAGGATTCGCGG 600
 Db 541 GTGGCGGTGGGCACTTATCTCAATGCAATAGCGTTCACTCTCACTAGGATTCGCGG 600
 QY 601 TTTCACCGAGTGGCGATCGCAGAGTTGATTCCTTTGAATTAAGGCGCTCGAGTTCCG 660
 Db 601 TTTCACCGAGTGGCGATCGCAGAGTTGATTCCTTTGAATTAAGGCGCTCGAGTTCCG 660
 QY 661 CCTTGGCTCTCTTTTTCAGCGTGAATTCCTGCGGAAAGAAAGAGATCCC 720
 Db 661 CCTTGGCTCTCTTTTTCAGCGTGAATTCCTGCGGAAAGAAAGAGATCCC 720
 QY 721 TTCTCTGCTGCTCGAGGTTTGAAGCTTCAACATGCTCTTGTGTAATTCAGGTCAGGC 780
 Db 721 TTCTCTGCTGCTCGAGGTTTGAAGCTTCAACATGCTCTTGTGTAATTCAGGTCAGGC 780
 QY 781 CCTATTGGCGGCGCTGATCTCTTGGGTCGTGACATCGGTAATTCCTTCTTGGG 840
 Db 781 CCTATTGGCGGCGCTGATCTCTTGGGTCGTGACATCGGTAATTCCTTCTTGGG 840
 QY 841 AAAGGCTCTAAATGACAACTGATTTCTCTGTAATTCCTTGTGTCAGATGTCGA 900
 Db 841 AAAGGCTCTAAATGACAACTGATTTCTCTGTAATTCCTTGTGTCAGATGTCGA 900
 QY 901 GTCAATTCTTTTGGCGCTCCGCGCGGTTCCGTTCTTAATCCCTTAAGCCCTCAAGGATCA 960
 Db 901 GTCAATTCTTTTGGCGCTCCGCGCGGTTCCGTTCTTAATCCCTTAAGCCCTCAAGGATCA 960
 QY 961 CAATTGTTGGGCAAAATGCGCATGTGATGCGCAGAGGATCTTGGCAATTTTGAACGCA 1020
 Db 961 CAATTGTTGGGCAAAATGCGCATGTGATGCGCAGAGGATCTTGGCAATTTTGAACGCA 1020
 QY 1021 TCACGCTTTTCGAGCAATGCGATGATGTAATCTGAAGATCTTAACCTTTGCTCAATGCGCTT 1080
 Db 1021 TCACGCTTTTCGAGCAATGCGATGATGTAATCTGAAGATCTTAACCTTTGCTCAATGCGCTT 1080
 QY 1081 GCGATTACGAGTGGCGGATCTTCTTGGCGGTCGACGCACTTTTGAAGGCTTGAAG 1140
 Db 1081 GCGATTACGAGTGGCGGATCTTCTTGGCGGTCGACGCACTTTTGAAGGCTTGAAG 1140
 QY 1141 GGCACCATGTTTGTGAGCTGTGATCTTTTCTAAACTGATTAATTAACAAAT 1200
 Db 1141 GGCACCATGTTTGTGAGCTGTGATCTTTTCTAAACTGATTAATTAACAAAT 1200
 QY 1201 CCGCATGCGCTCAATTTGAAGGAGATGCGGATTTTGAAGAACTTAAGAAAGCTTGAAG 1260
 Db 1201 CCGCATGCGCTCAATTTGAAGGAGATGCGGATTTTGAAGAACTTAAGAAAGCTTGAAG 1260
 QY 1261 CAGACAGCGCT 1271
 Db 1261 CAGACAGCGCT 1271

RESULT 5
 BD014991 1271 bp DNA linear PAT 27-AUG-2002
 LOCUS Nucleotide sequence encoding the discharge of branched amino acid,
 DEFINITION method of isolating the same and utilization thereof.
 ACCESSION BD014991.1 GI:22555798
 VERSION JP 2001169788-A/1.
 KEYWORDS unclassified
 SOURCE unclassified
 ORGANISM unclassified.
 1 (bases 1 to 1271)
 REFERENCE 1. Kerner, H., Eggering, L. and Pfeifferle, W.
 Nucleotide sequence encoding the discharge of branched amino acid,
 method of isolating the same and utilization thereof
 Patent: JP 2001169788-A 1 26-JUN-2001;
 DEGUSSA HUBLS AG, FORSCHUNGSZENTRUM JUELICH GMBH
 OS Corynebacterium glutamicum ATCC14752

PN JP 2001169788-A/1
 PD 26-JUN-2001 JP 2000324315
 PE 24-OCT-2000 DE 19951708.8
 PR 27-OCT-1999 DE 19951708.8
 PI NICOLE KERNERHUBER, HERMANN SAHN, LOTHAR EGGERING, WALTER PI
 PFEIFFERLE
 PC C12N15/09, C07K14/34, C12N1/21, C12P13/06, C12P13/08//
 PC C12N15/09, C12R1.15, (C12N1/21, C12R1.15), (C12P13/06, C12R1.15),
 PC C12N15/00, C12R1.15)
 CC brnF
 CC key
 FH Key
 FT Gene
 FT gene
 FEATURES Location/Qualifiers
 source 1..1271
 /organism="unclassified"
 /mol_type="genomic DNA"
 /db_xref="taxon:32644"

ORIGIN
 Query Match 100.0%; Score 1271; DB 6; Length 1271;
 Best Local Similarity 100.0%; Pred. No. 0; Indels 0; Gaps 0;
 Matches 1271; Conservative 0; Mismatches 0

QY 1 GCGCATCAATGGAATCTAGCTTCAATATTCACAAATAGCTAGTGAAGTGGCAAC 60
 Db 1 GCGCATCAATGGAATCTAGCTTCAATATTCACAAATAGCTAGTGAAGTGGCAAC 60
 QY 61 TGGCAAAAACCTACCCCGCAATTTGTGATGATTTGATGCAAAAACCGCAAGAT 120
 Db 61 TGGCAAAAACCTACCCCGCAATTTGTGATGATTTGATGCAAAAACCGCAAGAT 120
 QY 121 TCATTCAGGCTTGAAGTGTGCGCATCGCAAGGAGCGCTGGAACGAGATTAAGTTA 180
 Db 121 TCATTCAGGCTTGAAGTGTGCGCATCGCAAGGAGCGCTGGAACGAGATTAAGTTA 180
 QY 181 TGGCGGTACGAATGCGCAGGTCTAAACCTCCCTTGTGCAAGTTTGGCATGTA 240
 Db 181 TGGCGGTACGAATGCGCAGGTCTAAACCTCCCTTGTGCAAGTTTGGCATGTA 240
 QY 241 CCGGATGATGAGTGGTGGTCTCTGATTAATACGAGTGAAGTGGGCAAG 300
 Db 241 CCGGATGATGAGTGGTGGTCTCTGATTAATACGAGTGAAGTGGGCAAG 300
 QY 301 CCGCATGTTTTCGCGCTGATTTTTCGCGGCTCCACCGAAATCTGATCGCCCTCGT 360
 Db 301 CCGCATGTTTTCGCGCTGATTTTTCGCGGCTCCACCGAAATCTGATCGCCCTCGT 360
 QY 361 TGTGGGCGCAGCGCCCTGCGGCGCATCGGCTCAACCATTTGCTGTAATTCGCGCA 420
 Db 361 TGTGGGCGCAGCGCCCTGCGGCGCATCGGCTCAACCATTTGCTGTAATTCGCGCA 420
 QY 421 CGTATCTATGCGTTTCAATTCGCGCTGATGCTGCAAAAACCGCATTCGCTTTCTA 480
 Db 421 CGTATCTATGCGTTTCAATTCGCGCTGATGCTGCAAAAACCGCATTCGCTTTCTA 480
 QY 481 TTGGTTTTCGCGCTTATTCGACGAGGCTTACGAGTCACTGCGGCGCGCGAGGCTG 540
 Db 481 TTGGTTTTCGCGCTTATTCGACGAGGCTTACGAGTCACTGCGGCGCGCGAGGCTG 540
 QY 541 GTGGCGGTGGGCACTTATCTCAATGCAATAGCGTTCACTCTCACTAGGATTCGCGG 600
 Db 541 GTGGCGGTGGGCACTTATCTCAATGCAATAGCGTTCACTCTCACTAGGATTCGCGG 600
 QY 601 TTTCACCGAGTGGCGATCGCAGAGTTGATTCCTTTGAATTAAGGCGCTCGAGTTCCG 660
 Db 601 TTTCACCGAGTGGCGATCGCAGAGTTGATTCCTTTGAATTAAGGCGCTCGAGTTCCG 660
 QY 661 CCTTGGCTCTCTTTTTCAGCGTGAATTCCTGCGGAAAGAAAGAGATCCC 720
 Db 661 CCTTGGCTCTCTTTTTCAGCGTGAATTCCTGCGGAAAGAAAGAGATCCC 720

QY 421 CGTATTCATAGCGTTTCATTCGCCGCTGCGATGTCAGAAAACCCCATTTGCCGTTCTA 480
 DB 421 CGTATTCATAGCGTTTCATTCGCCGCTGCGATGTCAGAAAACCCCATTTGCCGTTCTA 480
 QY 481 TTCGGTTTCGGGCTTATTCAGAGAACCCGACATCTGGGGCGAGGCGGCGGCTG 540
 DB 481 TTCGGTTTCGGGCTTATTCAGAGAACCCGACATCTGGGGCGAGGCGGCGGCTG 540
 QY 541 GTCGGCGTGGCACTTATTCATATGCAATAGCAATAGCGTTTCACTCTCACTGGGTAATTCGGGG 600
 DB 541 GTCGGCGTGGCACTTATTCATATGCAATAGCAATAGCGTTTCACTCTCACTGGGTAATTCGGGG 600
 QY 601 TCTCACCGAGATGGCGATGCGAGATGTAATTCCTTTTGAATTAAGGCGCTCGAATTCGC 660
 DB 601 TCTCACCGAGATGGCGATGCGAGATGTAATTCCTTTTGAATTAAGGCGCTCGAATTCGC 660
 QY 661 CCTTTCCTCTCTCTTGTGACGCTGACTTGGATTCCTGCGAGCAAGAAAAGCAGATCCC 720
 DB 661 CCTTTCCTCTCTCTTGTGACGCTGACTTGGATTCCTGCGAGCAAGAAAAGCAGATCCC 720
 QY 721 TTCTCTGCTGCTGCGAGGTTTGAAGCTTCAACATGCTCTTGTGTAATTCAGGTCAGGC 780
 DB 721 TTCTCTGCTGCTGCGAGGTTTGAAGCTTCAACATGCTCTTGTGTAATTCAGGTCAGGC 780
 QY 781 CCTATTGCGGGCGCTGATCTTGGGTCGTTGACCATCCGTAATCTTCTTGGG 840
 DB 781 CCTATTGCGGGCGCTGATCTTGGGTCGTTGACCATCCGTAATCTTCTTGGG 840
 QY 841 AAGAGCTGCTAAATGACAACTGATTTCTCTGTAATTCCTTGTGTGCGAGTAATGTCGA 900
 DB 841 AAGAGCTGCTAAATGACAACTGATTTCTCTGTAATTCCTTGTGTGCGAGTAATGTCGA 900
 QY 901 GTCATTAATTTGGCGCTCCGCGGCTCCGTTCTTAAATCCTTAAGCCCTAGCATGATGCA 960
 DB 901 GTCATTAATTTGGCGCTCCGCGGCTCCGTTCTTAAATCCTTAAGCCCTAGCATGATGCA 960
 QY 961 CAATTGTGGGCAAAATGGCGATGTCATGTCAGCAGAGAACTCTTGCATTTTGAACCGCA 1020
 DB 961 CAATTGTGGGCAAAATGGCGATGTCATGTCAGCAGAGAACTCTTGCATTTTGAACCGCA 1020
 QY 1021 TCAAGCTTTGCGAGCATGTCATGTCAGCAGAGAACTCTTGCATTTTGAACCGCT 1080
 DB 1021 TCAAGCTTTGCGAGCATGTCATGTCAGCAGAGAACTCTTGCATTTTGAACCGCT 1080
 QY 1081 GCGATTAACAGTGGTGGCGCATCTTCTTGGCGGTCAGACGACCTTGTGAGCGTGGCGCT 1140
 DB 1081 GCGATTAACAGTGGTGGCGCATCTTCTTGGCGGTCAGACGACCTTGTGAGCGTGGCGCT 1140
 QY 1141 GGCACCATGCTTTTGTGAGTGTGATCTTTTCTAAACCTGCAATTAACCAAAAT 1200
 DB 1141 GGCACCATGCTTTTGTGAGTGTGATCTTTTCTAAACCTGCAATTAACCAAAAT 1200
 QY 1201 CCGCATGCGCTCAATTTGAAGGGGATGCGGATTTTGAAGAACTAGAAAAGGCTTAAG 1260
 DB 1201 CCGCATGCGCTCAATTTGAAGGGGATGCGGATTTTGAAGAACTAGAAAAGGCTTAAG 1260
 QY 1261 CAGACAGCGCT 1271
 DB 1261 CAGACAGCGCT 1271
 RESULT 4
 AX137714 1271 bp DNA linear PART 30-MAY-2001
 LOCUS AX137714
 DEFINITION Sequence 6 from Patent EP1096010.
 ACCESSION AX137714
 VERSION AX137714.1 GI:14273893
 KEYWORDS
 SOURCE Corynebacterium glutamicum
 ORGANISM Corynebacterium glutamicum
 Bacteria; Actinobacteria; Actinobacteriales; Actinomycetales;
 Corynebacteriaceae; Corynebacteriaceae; Corynebacterium.
 REFERENCE 1

AUTHORS Kernerkecht, N., Eggeling, L., Sahm, H. and Pfeifferle, W.
 TITLE Nucleotide sequences coding for branched-chain amino acids export
 JOURNAL proteins, method for isolating them and their use
 Patent: EP 1096010-A 6 02-MAY-2001;
 Degussa AG (DE); FORSCHUNGSZENTRUM JUELICH GMBH (DE)
 FEATURES
 source location/Qualifiers
 1..1271
 /organism="Corynebacterium glutamicum"
 /mol_type="unassigned DNA"
 /db_xref="taxon:1718"
 /note="ATCC13032"
 101..856
 /note="unassigned protein product; brnF"
 /codon_start=1
 /transl_table=1
 /protein_id="CAC39927.1"
 /db_xref="GI:14273894"
 /db_xref="REMBL:CAC39927"
 /translation="MKTQBIHSLSPSKALBDDPKYRREYIAQGLKSLAAGL
 GMPDIAFGLVIOYGYEWMAFLPSGLIFAGSTMLVIALVGAAPLALTL
 VNRHVFVAFSPFLHVKNPFIARFVSPALIDBAVTAAPAPGAMFLISQIAFH
 SYWRFQVAVAIABLIPEIKLEPALSIPVTLTLDSCRTKKQIPSLILGLSTFI
 ALVITPQALFPAHLIFLGILTRIFFLGKAA"
 853..1179
 /note="unassigned protein product; brnE"
 /codon_start=1
 /transl_table=1
 /protein_id="CAC39928.1"
 /db_xref="GI:14273895"
 /db_xref="REMBL:CAC39928"
 /translation="MTDFSCILLVAVACVITPALRANPFLIKPDRSQFVQKAM
 MWPAGIALITASTFRSNAIDLKTLFGLIAVAIVVAHLGGRLLSVAGTIVFV
 GLVNLFF"
 ORIGIN
 Query Match 100.0%; Score 1271; DB 6; Length 1271;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1271; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GCGCGATCAATGGAATGATCTTATGATATTTGACATATAGCTTATGAGTGGCGCAAC 60
 DB 1 GCGCGATCAATGGAATGATCTTATGATATTTGACATATAGCTTATGAGTGGCGCAAC 60
 QY 61 TGGCAACAACTACCCGCGAATTTGTGATGATTTGATGTCGAAACCGCAAGAT 120
 DB 61 TGGCAACAACTACCCGCGAATTTGTGATGATTTGATGTCGAAACCGCAAGAT 120
 QY 121 TCATTCAAGCTGGAAGTGTGCGCATCCAGAGCGCTGGAACCAAGATATAAGTTA 180
 DB 121 TCATTCAAGCTGGAAGTGTGCGCATCCAGAGCGCTGGAACCAAGATATAAGTTA 180
 QY 181 TGGGCGCTACGAAATTCGCGCAAGGCTTAACCTCCCTTGTGCAAGTTTGGCATGTA 240
 DB 181 TGGGCGCTACGAAATTCGCGCAAGGCTTAACCTCCCTTGTGCAAGTTTGGCATGTA 240
 QY 241 CCGCATTTGATGCTTGTGCTCTTGTATTAATCAATAGCGCTACGAATGATGGCGAGC 300
 DB 241 CCGCATTTGATGCTTGTGCTCTTGTATTAATCAATAGCGCTACGAATGATGGCGAGC 300
 QY 241 CCGCATTTGATGCTTGTGCTCTTGTATTAATCAATAGCGCTACGAATGATGGCGAGC 300
 DB 241 CCGCATTTGATGCTTGTGCTCTTGTATTAATCAATAGCGCTACGAATGATGGCGAGC 300
 QY 301 CCGCATTTTTCGCGGCTGATTTTCGCGGCTCCACCGGAATGCTGTCATCGCTCGT 360
 DB 301 CCGCATTTTTCGCGGCTGATTTTCGCGGCTCCACCGGAATGCTGTCATCGCTCGT 360
 QY 361 TGTGGGCGGAGCGCCCTGGGCGGCGCATGCGGCTACACCAATTTGATGGAATTCGCGCA 420
 DB 361 TGTGGGCGGAGCGCCCTGGGCGGCGCATGCGGCTACACCAATTTGATGGAATTCGCGCA 420
 QY 421 CGTATTCATAGCGTTTCATTCGCCGCTGCGATGTCAGAAAACCCCATTTGCCGTTCTA 480
 DB 421 CGTATTCATAGCGTTTCATTCGCCGCTGCGATGTCAGAAAACCCCATTTGCCGTTCTA 480
 QY 481 TTCGGTTTCGGGCTTATTCAGAGAACCCGACATCTGGGGCGAGGCGGCGGCTG 540

Db 301 CCCACTGTTTTCGGGCGTGAATTTTGGCGGGCTCCACCGAAATGCTGATCGCCCTCGT 360
 Qy 361 TGTGGGCGGAGGCGCCCTGGGCGCATCGCCGCTCACACATGCTGGTGAATCTCCGCA 420
 Db 361 TGTGGGCGGAGGCGCCCTGGGCGCATCGCCGCTCACACATGCTGGTGAATCTCCGCA 420
 Qy 421 CGTATTCATGAGGTTTTCATTCCTCCGTCATGCTGCAAAAACCCCATTTGCCGTTCTA 480
 Db 421 CGTATTCATGAGGTTTTCATTCCTCCGTCATGCTGCAAAAACCCCATTTGCCGTTCTA 480
 Qy 481 TTGGGTTTTCGGGCTTATTCAGAGAAAGCTTACGCACTGCTGCGGCGAGGCGCGAGGCTG 540
 Db 481 TTGGGTTTTCGGGCTTATTCAGAGAAAGCTTACGCACTGCTGCGGCGAGGCGCGAGGCTG 540
 Qy 541 GTCGGGCTGGGAGCTTATTCATGCAATGCAATAGCGTTTCACTGCTGCTGCTGCTGCTGCTG 600
 Db 541 GTCGGGCTGGGAGCTTATTCATGCAATGCAATAGCGTTTCACTGCTGCTGCTGCTGCTGCTG 600
 Qy 601 TCTCACCGGAGTGGCGATGCGAGAGTGAATTCCTTTGAATTAAGGCGCTCGAGTTGCG 660
 Db 601 TCTCACCGGAGTGGCGATGCGAGAGTGAATTCCTTTGAATTAAGGCGCTCGAGTTGCG 660
 Qy 661 CCTTTGCTCTCTCTTTGTCAGCGCTGACTTTGATTCCTGCGGAGGAGGAGGAGGAGGAGGAGG 720
 Db 661 CCTTTGCTCTCTCTTTGTCAGCGCTGACTTTGATTCCTGCGGAGGAGGAGGAGGAGGAGGAGG 720
 Qy 721 TTCTG 780
 Db 721 TTCTG 780
 Qy 781 CCTATTGGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840
 Db 781 CCTATTGGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840
 Qy 841 AAAGGCTGCTAAATGCACTGATTTCTCTGATTTCTCTGATTTCTCTGATTTCTCTGATTTCTCT 900
 Db 841 AAAGGCTGCTAAATGCACTGATTTCTCTGATTTCTCTGATTTCTCTGATTTCTCTGATTTCTCT 900
 Qy 901 GTCATTACTTTTGGGCTCCGGGCGGTTCCGTTCTTAATCTTAAGCCCTTAAGCTGAATCA 960
 Db 901 GTCATTACTTTTGGGCTCCGGGCGGTTCCGTTCTTAATCTTAAGCCCTTAAGCTGAATCA 960
 Qy 961 CAATTGTGGGCAAAATGGCGATGCTGATGCGAGAGAAATCTTGCCATTTTGACCGCA 1020
 Db 961 CAATTGTGGGCAAAATGGCGATGCTGATGCGAGAGAAATCTTGCCATTTTGACCGCA 1020
 Qy 1021 TCAACGTTTTCGAGCAATGCGATGATCTGAAGACTTAACTTTGCTCATTTGCCGT 1080
 Db 1021 TCAACGTTTTCGAGCAATGCGATGATCTGAAGACTTAACTTTGCTCATTTGCCGT 1080
 Qy 1081 GCGATTACAGTGTGGCGATCTTCTTGCGGCTGCAAGCACTTGTGAGCGCTGAGCGT 1140
 Db 1081 GCGATTACAGTGTGGCGATCTTCTTGCGGCTGCAAGCACTTGTGAGCGCTGAGCGT 1140
 Qy 1141 GGCACCATGTTTGTGAGTGTGAATCTTTTCTAAACTGATTAATAAATAAATAAATAAATAAATAA 1200
 Db 1141 GGCACCATGTTTGTGAGTGTGAATCTTTTCTAAACTGATTAATAAATAAATAAATAAATAAATAA 1200
 Qy 1201 CCGCATGCTCTCAATTTGAAGGGGATGCGGATTTTGAAGAACTTAAGAAAGGCTTAAG 1260
 Db 1201 CCGCATGCTCTCAATTTGAAGGGGATGCGGATTTTGAAGAACTTAAGAAAGGCTTAAG 1260
 Qy 1261 CAGACAGCGCT 1271
 Db 1261 CAGACAGCGCT 1271

RESULT 3
 AX137709 1271 bp DNA linear PAT 30-MAY-2001
 LOCUS AX137709
 DEFINITION Sequence 1 from Patent EP1096010.
 ACCSSION AX137709
 VERSION AX137709.1 GI:14273886

KEYWORDS
 SOURCE
 ORGANISM
 Corynebacterium glutamicum
 Corynebacterium glutamicum
 Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 Corynebacterineae; Corynebacteriaceae; Corynebacterium.
REFERENCE
 AUTHORS
 Kernerkecht, N., Eggeling, L., Sahm, H. and Pfeifferle, W.
 Nucleotide sequences coding for branched-chain amino acids export
 proteins, method for isolating them and their use
 Patent: EP 1096010-A 1 02-MAY-2001.
 Degussa AG (DE) ; FORSCHUNGSZENTRUM JUELICH GMBH (DE)
FEATURES
 source
 1. 1271
 Location/Qualifiers
 /organism="Corynebacterium glutamicum"
 /mol_type="unassigned DNA"
 /db_xref="taxon:1718"
 /note="ATCC14752"
 101..856
 /note="unnamed protein product; brnF"
 /codon_start=1
 /protein_id="CAC39923.1"
 /db_xref="GI:14273887"
 /db_xref="REMBL:CAC39923"
 /translation="MQRDRIHSSLSVSPSKAALBDDKGRYRYEIAQGLKSLAAGL
 GMPTPIGAFGLVIOGYEWMAPLFSGIIPAGSTMLVIALVVAAPGALATLL
 VNRHVFYAFSPFLHVKNPIDAFYVFLIDBAVTVTAIRPAGWAMBLISQIDFV
 SYVFGGLTGVLAELIPKIKGLFALSLFVTLTLDSCTRKQIPSLLAGLSFTI
 ALVIVGQALFALLIFLGLITRIFFLGMAK"
 853..1179
 /note="unnamed protein product; brnE"
 /codon_start=1
 /transl_table=11
 /protein_id="CAC39924.1"
 /db_xref="GI:14273888"
 /db_xref="REMBL:CAC39924"
 /translation="MTTPSCILLVAVCAVITFALRAVFLIKRLRSOPFYKAM
 WMPAGIATITASTFRSMALDKTLTFGLIAVAITVVAHLGGRRLLSVGACTIVF
 GLVNLV"
ORIGIN
 Query Match 100.0%; Score 1271; DB 6; Length 1271;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1271; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 1 GCGGATCAATGGAATGACTTACCTTCATATTTGCACAATTACCTTATGTTGAGGTGCGCAAC 60
 Qy 1 GCGGATCAATGGAATGACTTACCTTCATATTTGCACAATTACCTTATGTTGAGGTGCGCAAC 60
 Db 1 GCGGATCAATGGAATGACTTACCTTCATATTTGCACAATTACCTTATGTTGAGGTGCGCAAC 60
 Qy 61 TGGCAACAAACCTACCGGCAATTTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 120
 Db 61 TGGCAACAAACCTACCGGCAATTTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 120
 Qy 121 TCATTCAGCTGAGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
 Db 121 TCATTCAGCTGAGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
 Qy 121 TCATTCAGCTGAGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
 Db 121 TCATTCAGCTGAGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
 Qy 181 TGGGCGTACGAATGCGGCAAGGTCTAAACCTCCCTTGCTGAGAGTTTGGGCAATGTA 240
 Db 181 TGGGCGTACGAATGCGGCAAGGTCTAAACCTCCCTTGCTGAGAGTTTGGGCAATGTA 240
 Qy 181 TGGGCGTACGAATGCGGCAAGGTCTAAACCTCCCTTGCTGAGAGTTTGGGCAATGTA 240
 Db 181 TGGGCGTACGAATGCGGCAAGGTCTAAACCTCCCTTGCTGAGAGTTTGGGCAATGTA 240
 Qy 241 CCGCATTTGATTTGCGTTGGTCTTGTGTTATTCATTAACGCTACGAATGTTGGGCGAGC 300
 Db 241 CCGCATTTGATTTGCGTTGGTCTTGTGTTATTCATTAACGCTACGAATGTTGGGCGAGC 300
 Qy 241 CCGCATTTGATTTGCGTTGGTCTTGTGTTATTCATTAACGCTACGAATGTTGGGCGAGC 300
 Db 241 CCGCATTTGATTTGCGTTGGTCTTGTGTTATTCATTAACGCTACGAATGTTGGGCGAGC 300
 Qy 301 CCGCATGTTTTCGGGCTGATTTTTCGGGCTGCGGAGTCTGCGGAGTCTGCGGAGTCTGCGGAGTCT 360
 Db 301 CCGCATGTTTTCGGGCTGATTTTTCGGGCTGCGGAGTCTGCGGAGTCTGCGGAGTCTGCGGAGTCT 360
 Qy 361 TGTGGGCGGAGGCGCCCTGGGCGCATCGGCTCACCAATGCTGCTGTAATCTCCGCA 420
 Db 361 TGTGGGCGGAGGCGCCCTGGGCGCATCGGCTCACCAATGCTGCTGTAATCTCCGCA 420

FEATURES
source
Location/Qualifiers
1..1271
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN

Query Match 100.0%; Score 1271; DB 6; Length 1271;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1271; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 GCGGATCAATGGAATCTAGCTTCAATATATGCAATAGCTAGTGAAGTGGCAAC 60
1 GCGGATCAATGGAATCTAGCTTCAATATATGCAATAGCTAGTGAAGTGGCAAC 60
61 TGGCAACAAATCTACCCGGCAATTTGTGTATGATGTAGTGTGCAAAAAACGCAAGAT 120
61 TGGCAACAAATCTACCCGGCAATTTGTGTATGATGTAGTGTGCAAAAAACGCAAGAT 120
121 TCATTCAAGCTGAGAGGTGTGCGCATCCAAAGCAGCCCTGGAACAGATGATTAAGTTA 180
121 TCATTCAAGCTGAGAGGTGTGCGCATCCAAAGCAGCCCTGGAACAGATGATTAAGTTA 180
181 TCGGCGCTACGAAATCGCGCAAGGTCTAAAAACCTCCTTGTGCAAGTTTGGGCAATGTA 240
181 TCGGCGCTACGAAATCGCGCAAGGTCTAAAAACCTCCTTGTGCAAGTTTGGGCAATGTA 240
241 CCCGATGGTATTCGTTGGTCTCTTGTATTAATACGGCTACGAATGTGTGGCAGC 300
241 CCCGATGGTATTCGTTGGTCTCTTGTATTAATACGGCTACGAATGTGTGGCAGC 300
301 CCCGATGGTATTCGTTGGTCTCTTGTATTAATACGGCTACGAATGTGTGGCAGC 360
301 CCCGATGGTATTCGTTGGTCTCTTGTATTAATACGGCTACGAATGTGTGGCAGC 360
361 TGTGGGCGGAGCGCCCTGGCGGCGCATCGGCTCACACATTCCTGTGTAATTCGGCA 420
361 TGTGGGCGGAGCGCCCTGGCGGCGCATCGGCTCACACATTCCTGTGTAATTCGGCA 420
421 CGATATTCATGCGTTTCATTCCTCCGTCATGTGTGTCAAAAAACCCATTGCCGTTCTA 480
421 CGATATTCATGCGTTTCATTCCTCCGTCATGTGTGTCAAAAAACCCATTGCCGTTCTA 480
481 TTGGGTTTCCGCGCTATTCGACGAAGCTACGCAAGTACGCGCGCAGGCCGAGGCTG 540
481 TTGGGTTTCCGCGCTATTCGACGAAGCTACGCAAGTACGCGCGCAGGCCGAGGCTG 540
541 GTGCGCGTGGCGGCTTATTCGATGCAATAGCGTTTCACTCTCTAGTGGTATTCGCGG 600
541 GTGCGCGTGGCGGCTTATTCGATGCAATAGCGTTTCACTCTCTAGTGGTATTCGCGG 600
601 TCTCAACCGGAGTGGCGATCGCAAGTGTATTCCTTTGAAATTAAGGCGCTTCGAGTTCC 660
601 TCTCAACCGGAGTGGCGATCGCAAGTGTATTCCTTTGAAATTAAGGCGCTTCGAGTTCC 660
661 CCTTGTCTCTCTTTGTGCAAGCTGACTTTGATTTCTCGCGAAGCAAAAAAGCAGATCCC 720
661 CCTTGTCTCTCTTTGTGCAAGCTGACTTTGATTTCTCGCGAAGCAAAAAAGCAGATCCC 720
721 TTCTGTGCTGCTCGGAGTTGAGCTTCAATTCGCTTGTGTGTATTCAGGTCAAGC 780
721 TTCTGTGCTGCTCGGAGTTGAGCTTCAATTCGCTTGTGTGTATTCAGGTCAAGC 780
781 CCTATTTGCGGCGGCTGATCTTCTTGGGCTGTGTCATCCGATACCTTCTCTTGG 840
781 CCTATTTGCGGCGGCTGATCTTCTTGGGCTGTGTCATCCGATACCTTCTCTTGG 840
841 AAAGGCTCTAAATGACAACTGATTTCTCTGTAATCTCTTGTGTGTGCGAGTATGCA 900
841 AAAGGCTCTAAATGACAACTGATTTCTCTGTAATCTCTTGTGTGTGCGAGTATGCA 900
901 GTCAATTAATTTGCGCGCGGCGGCTGCTTAAATCCTTAAGCCCTTACGTAATCA 960
901 GTCAATTAATTTGCGCGCGGCGGCTGCTTAAATCCTTAAGCCCTTACGTAATCA 960

ORIGIN

Query Match 100.0%; Score 1271; DB 6; Length 1271;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1271; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

961 CAATTTGTGGCAAAATGGCGGATGTGATGCCAGCAGGAATCCTTGCCATTTTGACCGCA 1020
961 CAATTTGTGGCAAAATGGCGGATGTGATGCCAGCAGGAATCCTTGCCATTTTGACCGCA 1020
1021 TCAAGTTTGGCAGCAATGCGATAGATCTGAGACTCTAACTTTGTCTCATTTGCCGTT 1080
1021 TCAAGTTTGGCAGCAATGCGATAGATCTGAGACTCTAACTTTGTCTCATTTGCCGTT 1080
1081 GCGATTCAGTGTGGCGCATCTTCTTGGCGGTGAGACGACCTTTGTGACGTTGGCGCT 1140
1081 GCGATTCAGTGTGGCGCATCTTCTTGGCGGTGAGACGACCTTTGTGACGTTGGCGCT 1140
1141 GGCACATCGTTTGTGTGAGCTGGTGAATCTTTCTAAACTGCAATAAACAATAAT 1200
1141 GGCACATCGTTTGTGTGAGCTGGTGAATCTTTCTAAACTGCAATAAACAATAAT 1200
1201 CCGCATGCTCTCAATTTGAAGGGGATGGGATTTTAAAGGACTTGAAAAGGCTTTAG 1260
1201 CCGCATGCTCTCAATTTGAAGGGGATGGGATTTTAAAGGACTTGAAAAGGCTTTAG 1260
1261 CAGACAGCGCT 1271
1261 CAGACAGCGCT 1271

RESULT 2
AR391956 1271 bp DNA linear PAT 18-DEC-2003
LOCUS AR391956
DEFINITION Sequence 6 from patent US 6613545.
ACCESSION AR391956
VERSION AR391956.1 GI:40115727
KEYWORDS
SOURCE
ORGANISM
Unknown.
Unclassified.
REFERENCE
1 (bases 1 to 1271)
Kemernecht,N., Salm,H., Eggeling,L. and Pfeifferle,W.
Nucleotide sequences coding for the export of branched chain amino acids, process for the isolation thereof and use thereof
JOURNAL
Patent: US 6613545-A 6 02-SEP-2003;
FEATURES
source
Location/Qualifiers
1..1271
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN

Query Match 100.0%; Score 1271; DB 6; Length 1271;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1271; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 GCGGATCAATGGAATCTAGCTTCAATATATGCAATAGCTAGTGAAGTGGCAAC 60
1 GCGGATCAATGGAATCTAGCTTCAATATATGCAATAGCTAGTGAAGTGGCAAC 60
61 TGGCAACAAATCTACCCGGCAATTTGTGTATGATGTAGTGTGCAAAAAACGCAAGAT 120
61 TGGCAACAAATCTACCCGGCAATTTGTGTATGATGTAGTGTGCAAAAAACGCAAGAT 120
121 TCATTCAAGCTGAGAGGTGTGCGCATCCAAAGCAGCCCTGGAACAGATGATTAAGTTA 180
121 TCATTCAAGCTGAGAGGTGTGCGCATCCAAAGCAGCCCTGGAACAGATGATTAAGTTA 180
181 TCGGCGCTACGAAATCGCGCAAGGTCTAAAAACCTCCTTGTGCAAGTTTGGGCAATGTA 240
181 TCGGCGCTACGAAATCGCGCAAGGTCTAAAAACCTCCTTGTGCAAGTTTGGGCAATGTA 240
241 CCCGATGGTATTCGTTGGTCTCTTGTATTAATACGGCTACGAATGTGTGGCAGC 300
241 CCCGATGGTATTCGTTGGTCTCTTGTATTAATACGGCTACGAATGTGTGGCAGC 300
301 CCCGATGGTATTCGTTGGTCTCTTGTATTAATACGGCTACGAATGTGTGGCAGC 360
301 CCCGATGGTATTCGTTGGTCTCTTGTATTAATACGGCTACGAATGTGTGGCAGC 360

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 19, 2004, 10:57:18 ; Search time 5387 Seconds

(without alignments)
10226.291 Million cell updates/sec

Title: US-10-608-504-6

Perfect score: 1271

Sequence: 1 gcgcgcacatgcaatctagct.....agccttaagcagacagcgcct 1271

Scoring table: IDENTITY NTC

Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 2167151695 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: gb_ba:*
2: gb_hcg:*
3: gb_in:*
4: gb_cm:*
5: gb_ov:*
6: gb_ph:*
7: gb_pl:*
8: gb_pr:*
9: gb_ro:*
10: gb_sts:*
11: gb_sy:*
12: gb_un:*
13: gb_vl:*
14: gb_wi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_or:*
21: em_cr:*
22: em_cv:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vl:*
30: em_hcg_hum:*
31: em_hcg_inv:*
32: em_hcg_other:*
33: em_hcg_mus:*
34: em_hcg_pin:*
35: em_hcg_rnd:*
36: em_hcg_mam:*
37: em_hcg_vrt:*
38: em_sy:*
39: em_hcg_hum:*
40: em_hcg_mus:*
41: em_hcg_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	1271	100.0	1271	6 AR391953
2	1271	100.0	1271	6 AR391956
3	1271	100.0	1271	6 AX137709
4	1271	100.0	1271	6 AX137714
5	1271	100.0	1271	6 BD014991
6	1271	100.0	1271	6 BD014992
7	1271	100.0	340000	6 AP005274
8	1271	100.0	349980	6 AP005274
9	1260	99.1	2105	1 AF454053
10	753	59.2	753	6 AR391954
11	753	59.2	753	6 AX120373
12	753	59.2	753	6 AX137710
13	753	59.2	753	6 BD014992
14	753	59.2	753	6 BD162490
15	365.6	28.8	397	6 AX137083
16	365.6	28.8	778	6 AX137085
17	324	25.5	324	6 AR391955
18	324	25.5	324	6 AX120374
19	324	25.5	324	6 AX137712
20	324	25.5	324	6 BD014993
21	324	25.5	324	6 BD162491
22	185	14.6	715	6 AX137077
23	185	14.6	715	6 BD013995
24	149.6	11.8	87340	1 AP005224
25	132.8	10.4	53793	1 BX248361
26	127.4	10.0	711	6 AX123492
27	127.4	10.0	711	6 BD162609
28	127.4	10.0	309400	6 AX127153
29	127.4	10.0	325651	1 AP005283
30	59.8	4.7	349926	1 BX571660
31	58.6	4.6	756	1 AR366287
32	57.2	4.5	9888	1 AE011050
33	54.4	4.3	856	1 HSA352448
34	52.4	4.1	7218	6 166494
35	51.8	4.1	11025	1 AE013585
36	51.2	4.0	125020	9 AF429315
37	51	4.0	308015	1 AE016783
38	48.2	3.8	2000	6 AX655393
39	48	3.8	10732	6 E32986
40	46.6	3.7	3619	1 ECH410307
41	46.6	3.7	10733	1 AE013695
42	46.6	3.7	220050	1 AJ414156
43	46.6	3.6	215050	1 AL646084
44	46	3.6	20342	1 AE008888
45	45.8	3.6	20342	1 AE008888

ALIGNMENTS

RESULT 1
AR391953
LOCUS AR391953 1271 bp DNA linear PAT 18-DEC-2003
DEFINITION Sequence 1 from patent US 6613545.
ACCESSION AR391953
VERSION AR391953.1 GI:40115724
KEYWORDS
ORGANISM Unknown.
SOURCE Unknown.
REFERENCE 1 (bases 1 to 1271)
AUTHORS Kennetknecht,N., Salm,H., Eggeling,L. and Pfeifferle,M.
TITLE Nucleotide sequences coding for the export of branched chain amino acids, process for the isolation thereof and use thereof
JOURNAL Patent: US 6613545-A 1 02-SEP-2003;

RESULT 15

O8TKG9

ID O8TKG9 PRELIMINARY; PRT: 113 AA.

AC O8TKG9;

DT 01-JUN-2002 (TrEMBLrel. 21, Created)

DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

DE Branched chain amino acid transport protein Azid.

GN MA3438.

OS Methanosarcina acetivorans.

OC Archaea; Euryarchaeota; Buryarchaeota orders incertae sedis;

OC Methanosarcinales; Methanosarcinaceae; Methanosarcina.

OX NCBI_TaxID=2214;

RN (1)

RP SEQUENCE FROM N.A.

RC STRAIN=C2A / ATCC 35395 / DSM 2834;

RX MEDLINE=21929760; PubMed=1932238;

RA Galagan J.E., Nusbaum C., Roy A., Endrizzi M.G., Macdonald P.,

RA Fitzhugh W., Calvo S., Engels R., Smitrov S., Arnoor D., Brown A.,

RA Allen N., Naylor J., Strange-Thomann N., Dezelicno K., Johnson R.,

RA Linton L., McEwan P., McKernan K., Talamas J., Tittell A., Ye W.,

RA Zimmer A., Barber R.D., Cam I., Graham D.E., Grahame D.A., Guss A.M.,

RA Hedderich R., Ingram-Smith C., Kuetner H.C., Krzycki J.A.,

RA Leigh J.A., Li W., Liu J., Mukhopadhyay B., Reeve J.N., Smith K.,

RA Springer T.A., Umayam L.A., White O., White R.H., de Macario E.C.,

RA Ferry J.G., Jarrell K.F., Jing H., Macario A.J.L., Paulsen I.,

RA Pritchett M., Sowers K.R., Swanson R.V., Zinder S.H., Lander E.,

RA Metcalf W.W., Birren B.;

RT "The genome of Methanosarcina acetivorans reveals extensive metabolic

RL and physiological diversity.";

RL Genome Res. 12:532-542(2002).

DR EMBL: AE011050; AAM06805.1; -

DR InterPro: IPR008407; Azid.

DR Pfam: PF05437; Azid; 1.

KW Complete proteome.

SQ SEQUENCE 113 AA; 12626 MW; 90CBE128F93D7D CRC64;

Query Match 15.6%; Score 82; DB 17; Length 113;

Best Local Similarity 28.6%; Pred. No. 1;

Matches 32; Conservative 20; Mismatches 54; Indels 6; Gaps 4;

QY 1 MTTDESCILIVAVCAVITPALRAVPFLIKPRESQFVGKAMMMFAGILAITASTFR 60
 DB 1 MMEPTLOMATIVIALATFATFVLPPLCGSRREPPAMLSITEKNLPMILLLVICYCK 60
 QY 61 SNAIDKLTLPGLIAV-AITVAHL-LGGRTLLSYGAGTIVGVV--NLF 108
 DB 61 D--VQWFLAPYGIPELFTIGIVAGLHMKRNMLSTFAGTGLVMLVQFNVF 110

Search completed: April 16, 2004, 06:55:52
 Job time : 23.3593 secs

DR Pfam: PF00211; guanylate_cyc; 2.
 DR SMART: SM00044; CYC; 2.
 DR PROSITE: PS00452; GUANYLATE_CYCLASES_1; 2.
 DR PROSITE: PS50125; GUANYLATE_CYCLASES_2; 2.
 KW Lyase.
 FT NON_TER 1
 FT NON_TER 952
 SQ SEQUENCE 952 AA; 106967 MW; C7F30C7A67E6C5B8 CRC64;

Query Match
 Best Local Similarity 31.4%; Pred. No. 3.5;
 Matches 33; Conservative 19; Mismatches 31; Indels 22; Gaps 7;

QY 18 ITFALRAVPLIKP---LRESQF---VGRMMMPAGILALITASTFSSNAI---DLK 67
 DB 31 LTFALRCFPALGAPALQEAQWARTVAEQWQ-----LLVTFVSVALPVSIL 83

QY 68 TLTFGLIAVA--ITVAHLGGRTLL--SVGAGTIVFGLVNL 108
 DB 84 AIGFGLVVAASHLVATLVPAKPRPRLWRTLGANALLFGV-VNMY 127

RESULT 9
 Q8NFMS PRELIMINARY; PRT; 301 AA.
 ID Q8NFMS
 AC Q8NFMS
 DT 01-OCT-2002 (TREMBLrel. 22, Created)
 DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)
 DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
 DE Adenylate cyclase type I (EC 4.6.1.1) (Fragment).
 GN AOCY1
 OS Homo sapiens (Human)
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OC NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22391367; PubMed=12503609;
 RA Ludwig M.G., Seuwen K.;
 RT "Characterization of the human adenylate cyclase gene family: cDNA,
 RT gene structure, and tissue distribution of the nine isoforms."
 RL J. Recept. Signal Transduct. Res. 22:79-110(2002).
 DR EMBL: AF497515; AAM94372.1;
 DR GO: GO:0008294; F:calcium/calmodulin-responsive adenylate cyc. .; IEA.
 DR GO: GO:0016829; F:lyase activity; IEA.
 KM Lyase
 FT NON_TER 301
 FT NON_TER 301
 SQ SEQUENCE 301 AA; 32062 MW; 6AB95F1B98F466C6 CRC64;

Query Match
 Best Local Similarity 31.4%; Pred. No. 1.4;
 Matches 33; Conservative 18; Mismatches 44; Indels 10; Gaps 5;

QY 12 VAVCAVITFARAVPLIKPLRESQFVGKAMMPGILALITASTFSSNAI---DLK 67
 DB 125 LALPFLSLFALCFALGAPARSAAGAPARABOGWOLIV-TFVSALLPVSIL 133

QY 68 TLTFGLIAVA--ITVAHLGGRTLL--SVGAGTIVFGLVNL 108
 DB 184 AIGFGLVVAASHLVATLVPAKPRPRLWRTLGANALLFGV-VNMY 227

RESULT 10
 Q92BR3 PRELIMINARY; PRT; 108 AA.
 ID Q92BR3
 AC Q92BR3
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
 DE Hypothetical protein lin1481.
 GN LIN1481.
 OS Listeria innocua.
 OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.

OX NCBI_TaxID=1642;
 RN [1]
 RC SEQUENCE FROM N.A.
 RC STRAIN=CIP 1162 / Serovar 6a;
 RX MEDLINE=21537279; PubMed=11679669;
 RA Glaeser P., Frangeul L., Buchrieser C., Rusnlok C., Amend A.,
 RA Baquero F., Berche P., Biochecker H., Brandt P., Chakraborty T.,
 RA Charbit A., Chetouani F., Couve E., de Daruvar A., Deloux P.,
 RA Domant E., Dominguez-Bernal G., Duchaud E., Durant L., Dussauget O.,
 RA Entian K.-D., Feih W., Garcia-del Portillo F., Garrido P., Jackson D.,
 RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Haut U., Jackson D.,
 RA Jones L.-M., Kaefer U., Kreft U., Kuhn M., Kunst F., Kuepfer G.,
 RA Madueno E., Matounam A., Mata Vicente J., Ng E., Nedjati H.,
 RA Nordstedt G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,
 RA Remmel B., Rose M., Schluter T., Simoes N., Tierrez A.,
 RA Vazquez-Boland J.-A., Vose H., Weiland J., Cossart P.,
 RT "Comparative genomics of Listeria species."
 RL Science 294:849-852(2001).
 DR EMBL: AL596168; CAC96712.1; .
 DR PIR: AH1617; AH1617.
 DR Listlist: LIN01481; .
 KM Hypothetical protein; Complete proteome.
 SQ SEQUENCE 108 AA; 12014 MW; 89B14D6C6B8E083 CRC64;

Query Match
 Best Local Similarity 24.8%; Pred. No. 0.54;
 Matches 25; Conservative 24; Mismatches 43; Indels 9; Gaps 3;

QY 9 LUVAVCAVITFARAVPLIKPLRESQFVGKAMMPAGILALITASTFSSNAIDLK 67
 DB 8 LVLVGGCLVFIFRVLPFVRKLPSPVIRLSVLPCLITLAFVGSLLITNENSP 67

QY 68 TLTFGLIAVA--ITVAHLGGRTLL--SVGAGTIVFGLVNL 107
 DB 68 SINENLALTPITIT-----ALTFKMLMIVYIGISM 101

RESULT 11
 Q7VI97 PRELIMINARY; PRT; 113 AA.
 ID Q7VI97
 AC Q7VI97
 DT 01-OCT-2003 (TREMBLrel. 25, Created)
 DT 01-OCT-2003 (TREMBLrel. 25, Last sequence update)
 DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
 DE Hypothetical protein.
 GN HH0711.
 OS Helicobacter hepaticus.
 OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
 OC Helicobacteraceae; Helicobacter.
 OC NCBI_TaxID=32025;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 51449 / 3B1;
 RX MEDLINE=22709201; PubMed=12810954;
 RX Suerbaum S., Josephans C., Stenzenbach T., Drescher B., Brandt P.,
 RA Bell M., Droege M., Partmann B., Fischer H.-P., Ge Z., Hoerster A.,
 RA Holland R., Klein K., Koening J., Macko L., Mendz G.L., Nyakatura G.,
 RA Schauer D.B., Shen Z., Weber J., Froesch M., Fox J.G.;
 RT "The complete genome sequence of the carcinogenic bacterium
 RT Helicobacter hepaticus."
 RL Proc. Natl. Acad. Sci. U.S.A. 100:7901-7906(2003).
 DR EMBL: AE017146; AAP77308.1; .
 KM Hypothetical protein; Complete proteome.
 SQ SEQUENCE 113 AA; 12511 MW; 6D613F77CBB6C90F CRC64;

Query Match
 Best Local Similarity 27.0%; Pred. No. 0.57;
 Matches 27; Conservative 27; Mismatches 39; Indels 7; Gaps 4;

QY 9 LUVAVCAVITFARAVPLIKPLRESQFVGKAMMPAGILALITASTFSSNAIDLK 68
 DB 9 VLIALIALNTLSRLPPIIFAKSTPT-FIVSLGKVPESALIMLIYCLKD--MDLKH 65

RT avermitilis: deducing the ability of producing secondary
 metabolites";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
 RX MEDLINE=22608306; PubMed=12692562;
 RA Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,
 Sakaki Y., Hattori M., Omura S.;
 RT "Complete genome sequence and comparative analysis of the industrial
 microorganism *Streptomyces avermitilis*,"
 RL Nat. Biotechnol. 21:526-531(2003).
 DR EMBL; AF005031; BAC70208.1; -
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 102 AA; 10210 MW; 1D803E3D1A40563B CRC64;

Query Match 17.1%; Score 90; DB 16; Length 102;
 Best Local Similarity 29.7%; Pred. No. 0.14;
 Matches 30; Conservative 18; Mismatches 43; Indels 10; Gaps 3;
 QY 8 ILVVAVCAVITPALRAVPFLIKPLRESQFVKMMMPAGIATLTA-STP---RSNA 63
 DB 3 IMAIAATVAGCVAVGLIGLIVPAGVLERPLVRLLALPVALTAALTAOCTFADGRTLV 62
 QY 64 IDLKTLTFLGLIAVITVAHLGGRRITLISVAGTIVFGL 104
 DB 63 LDAXAGLAIAAAVAVLRAFP-----LLVAAAVVTVAGV 97

RESULT 6
 Q8G3T6 PRELIMINARY; PRT; 110 AA.

AC Q8G3T6;
 DT 01-MAR-2003 (TREMBlrel. 23, Created)
 DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
 DE Branched-chain amino acid permease.
 GN ALZD OR BL1668.
 OS Bifidobacterium longum.
 OC Bacteria; Actinobacteria; Actinobacteriales; Bifidobacteriales;
 OC Bifidobacteriaceae; Bifidobacterium.
 OX NCBI_TaxID=216816;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NCC 2705;
 RX MEDLINE=22284977; PubMed=12381787;
 RA Scheil M.A., Karmirantzou M., Snel B., Vilarova D., Berger B.,
 Pessi G., Zehren M.-C., Desiere F., Bork P., Delley M.,
 RT "The genome sequence of *Bifidobacterium longum* reflects its adaptation
 to the human gastrointestinal tract,"
 RT Proc. Natl. Acad. Sci. U.S.A. 99:14422-14427(2002).
 RL EMBL; AE014601; AAN25455.1; -
 DR InterPro; IPR008407; AZID.
 DR Pfam; PF05437; AZID; 1.
 KW Complete proteome.
 SQ SEQUENCE 110 AA; 12265 MW; 52FB38C7F95274CC CRC64;

Query Match 16.8%; Score 88.5; DB 16; Length 110;
 Best Local Similarity 30.7%; Pred. No. 0.22;
 Matches 35; Conservative 19; Mismatches 45; Indels 15; Gaps 6;
 QY 1 MTTFSCILVVAVCAVITPALRAVPFLI---KPLRESQFVKMMMPAGIATLTA 56
 DB 3 MTIMOGAIVTVVLG--TWCTRFLLPFLVPESKOPRIIEYFGV--LPYMTGLLV 57
 QY 57 STRSNADIKTLTFLG---LIIVAVITVAHLGGRRITLISVAGTIVFGLVNL 107
 DB 58 YALNNTPI--LTGSHGLPELIACIVYLHL-VKGNMLLSIAGTIVYMLLVOL 108

RESULT 7
 Q8PS24

ID Q8PS24 PRELIMINARY; PRT; 112 AA.
 AC Q8PS24;
 DT 01-OCT-2002 (TREMBlrel. 22, Created)
 DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
 DE Branched-chain amino acid transport protein.
 GN ALZD OR MM3263.
 OS Methanosarcina mazei (Methanosarcina frisia).
 OC Archaea; Euryarchaeota; Euryarchaeota orders incertae sedis;
 OC Methanosarcinales; Methanosarcinaceae; Methanosarcina.
 OX NCBI_TaxID=2209;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Gael / G01 / ATCC BAA-199 / DSM 3647 / OCM 88;
 RX MEDLINE=22120827; PubMed=12125824;
 RA Deppenmeier U., Johann A., Hartsch T., Merkl R., Schmitz R.A.,
 Martinez-Arias R., Henne A., Wieser A., Baumer S., Jacobl C.,
 Brueggemann H., Llenard T., Christmann A., Boemcke M., Steckel S., P.,
 RA Bhattacharyya A., Lykidis A., Overbeek R., Klenk H.-P., Gunsalus R.P.,
 RA Fritz H.-J., Gottschalk G.;
 RT "The genome of *Methanosarcina mazei*: evidence for lateral gene
 transfer between bacteria and archaea,"
 RL J. Mol. Microbiol. Biotechnol. 4:453-461(2002).
 DR EMBL; AE013585; AAM32959.1; -
 DR InterPro; IPR008407; AZID.
 DR Pfam; PF05437; AZID; 1.
 KW Complete proteome.
 SQ SEQUENCE 112 AA; 12433 MW; 33D73C78B8DDA65C CRC64;

Query Match 16.5%; Score 87; DB 17; Length 112;
 Best Local Similarity 29.4%; Pred. No. 0.32;
 Matches 30; Conservative 19; Mismatches 49; Indels 4; Gaps 3;
 QY 8 ILVVAVCAVITPALRAVPFLIKPLRESQFVKMMMPAGIATLTA-STP---RSNA 67
 DB 7 MLVIAIATATFTRTLPLPLCGSRPPMLSTVENLPEMLLLIVCLND--VQWF 64
 QY 68 TLFGLIAV-AITVAHL-LGRRITLISVAGTIVFGLVNL 107
 DB 65 SAPYGFPELFTIGVAGLHFVKMMMLSPAGTIGIYVALVOL 106

RESULT 8

ID Q88444 PRELIMINARY; PRT; 952 AA.
 AC Q88444;
 DT 01-NOV-1998 (TREMBlrel. 08, Created)
 DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
 DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
 DE Adenylyl cyclase type I (EC 4.6.1.1) (Fragment).
 GN ADCT1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Brain;
 RX MEDLINE=98324784; PubMed=9662407;
 RA Abdel-Majid R.M., Leong W.L., Schalkwyk L.C., Smallman D.S.,
 Wong S.T., Storm D.R., Fine A., Dobson M.J., Guernsey D.L.,
 RA Neumann P.E.;
 RT "Loss of adenylyl cyclase I activity disrupts patterning of mouse
 somato sensory cortex,"
 RT Nat. Genet. 18:289-291(1998).
 RL EMBL; AF053980; AAC29478.1; -
 DR HSSP; P26769; IAB8.
 DR MGD; MGI:99677; Adcy1.
 DR GO; GO:0008294; F:calcium/calmodulin-responsive adenylylate cyc. .; IEA.
 DR GO; GO:0004383; F:guanylate cyclase activity; IEA.
 DR GO; GO:0016829; F:lyase activity; IEA.
 DR GO; GO:0007242; P:intracellular signaling cascade; IEA.
 DR InterPro; IPR01054; G_cyclase.

Db 1 MTPTSCILLVAVCAVITPALRAVPLILKPRESOFGKMMMPAGIILITASTFR 60
 61 SNAIDKTLTFFGLIAVAITVAHLLGGRRTLLSVGAGTIVFVGLVNL 108
 61 SNAIDKTLTFFGLIAVAITVAHLLGGRRTLLSVGAGTIVFVGLVNL 108

RESULT 2

Q8NL67 PRELIMINARY; PRT; 115 AA.
 AC Q8NL67;
 DT 01-OCT-2002 (TREMBlrel. 22, Created)
 DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
 DE Hypothetical membrane protein Cg13082.
 GN Cg13082.
 OS Corynebacterium glutamicum (Brevibacterium flavum).
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacteriaceae; Corynebacteriaceae; Corynebacterium.
 OX NCBI_TaxID=1718;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;
 RA Nakagawa S.;
 RT "Complete genomic sequence of Corynebacterium glutamicum ATCC 13032."
 RL Submitted (May-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF005283; BAC00476.1; -
 DR InterPro; IPR008407; AZID.
 DR Pfam; PF05437; AZID; 1.
 KW Complete proteome; Hypothetical protein.
 SQ SEQUENCE 115 AA; 12018 MW; D18F7D197347F7D CRC64;

Query Match 24.7%; Score 130; DB 16; Length 115;
 Best Local Similarity 35.0%; Pred. No. 1.6e-05;
 Matches 36; Conservative 16; Mismatches 47; Indels 4; Gaps 2;

QY 8 ILVVAV---CAVITPALRAVPLILKPRESOFGKMMMPAGIILITASTFRSNAI 64
 12 LNVAAVLTIPALITLLRIFPFAMKRVNSNOLMGVLRITPVGVVTVLITLFGQVS 71
 QY 65 DLKTLTFFGLIAVAITVAHLLGGRRTLLSVGAGTIVFVGLVNL 107
 72 APGVGASLIIVAFALHMLKQSAAGLIVG-GLIIVMPLVNV 113

RESULT 3

Q8FSW9 PRELIMINARY; PRT; 116 AA.
 AC Q8FSW9;
 DT 01-MAR-2003 (TREMBlrel. 23, Created)
 DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
 DE Conserved hypothetical protein.
 GN CE2925.
 OS Corynebacterium efficiens.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacteriaceae; Corynebacteriaceae; Corynebacterium.
 OX NCBI_TaxID=152794;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=YS-314 / AJ 12310 / DSM 44549 / JCM 11189;
 RA Kawarabayashi Y., Yamazaki J., Hino Y., Kikuchi H., Nakamura Y.,
 RA Ikeo K., Suzuki M., Mashima J., Itoh T., Yamagishi A., Nishio Y.,
 RA Ueda Y., Sugimoto S.;
 RT "The entire genomic sequence of Corynebacterium efficiens YS-314."
 RL Submitted (May-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF005224; BAC19735.1; -
 DR InterPro; IPR008407; AZID.
 DR Pfam; PF05437; AZID; 1.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 116 AA; 12255 MW; 31B0AAE454103417 CRC64;

Query Match 24.4%; Score 128.5; DB 16; Length 116;
 Best Local Similarity 30.7%; Pred. No. 2.2e-05;
 Matches 31; Conservative 27; Mismatches 40; Indels 3; Gaps 2;

QY 8 ILVVAVCAVITPALRAVPLILKPRESOFGKMMMPAGIILITASTFRSNAIDK 67
 15 VAVLIPVAILITLRLFPFVAMRGVKNQMLGVLSRMTIPGVSVLITLFTSTGSPG 74
 QY 68 TLTFGLIAVAITVAHLLGGRRTLLSVGAGTIVFVGLVNL 107
 75 GIWASLIIVGITLILH--WMRSAGLIVGTIVAVTVLVNV 113

RESULT 4

Q9CF69 PRELIMINARY; PRT; 108 AA.
 AC Q9CF69;
 DT 01-JUN-2001 (TREMBlrel. 17, Created)
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
 DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
 DE Unknown protein.
 GN Y0RC OR L1612.
 OS Lactococcus lactis (subsp. lactis) (Streptococcus lactis).
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.
 OX NCBI_TaxID=1360;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=IL1403;
 RX MEDLINE=2125186; PubMed=11337471;
 RA Bofurin A., Wincker P., Manger S., Jallion O., Malarme K.,
 RA Weisenbach J., Ehrlich S.D., Sorokin A.;
 RT "The complete genome sequence of the lactic acid bacterium Lactococcus
 RT lactis ssp. lactis IL1403."
 RL Genome Res. 11:731-753(2001).
 DR EMBL; AE006391; AAK05710.1; -
 DR PIR; D86826; D86826.
 KW Complete proteome.
 SQ SEQUENCE 108 AA; 12024 MW; 9EE4838C1DFCF972 CRC64;

Query Match 18.0%; Score 94.5; DB 16; Length 108;
 Best Local Similarity 29.9%; Pred. No. 0.054;
 Matches 32; Conservative 24; Mismatches 42; Indels 9; Gaps 4;

QY 3 TDPSCTLLVAVCAVITPALRAVPLILKPRESOFGKMMMPAGIILITAST-FRS 61
 2 SSFESITLIGCAITWISRLVPLILKQSLPQIVVEYISFVVMISMLWISNLFIQ 61
 Db 62 NAIDKTLTFF-GLIIVAVHLLGGRRTLLSVGAGTIVFVGLVNL 107
 62 HLGHLPSVNMNMLASIPVLAALD--TKNLL----VIVLVGFSL 101

RESULT 5

Q82XA6 PRELIMINARY; PRT; 102 AA.
 AC Q82XA6;
 DT 01-JUN-2003 (TREMBlrel. 24, Created)
 DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
 DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
 DE Hypothetical protein.
 GN SAV2497.
 OS Streptomyces avermitilis.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Streptomycetaceae; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxID=33903;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
 RX MEDLINE=21477403; PubMed=11572948;
 RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
 RA Shinoue M., Ikeda H., Horikawa H., Nakazawa H., Osone T.,
 RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.;
 RT "Genome sequence of an industrial microorganism Streptomyces

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 16, 2004, 06:51:17 ; Search time 21.3593 Seconds
(without alignments)
1595.367 Million cell updates/sec

Title: US-10-608-504-5

Perfect score: 526

Sequence: 1 MTTDPSCLLVAVACAVITF.....RTLLSVGAGTIVFGLVLP 108

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

SPTRMBL_25:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriophage:*
17: sp_archaeal:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	526	100.0	108	16	Q8NTP7 corynebacte
2	130	24.7	115	16	Q8NTL67 corynebacte
3	128.5	24.4	116	16	Q8NTSW9 corynebacte
4	94.5	18.0	108	16	Q8CF69 lactococcus
5	90	17.1	102	16	Q82KX6 streptomyces
6	88.5	16.8	110	16	Q8G3T6 bifidobacte
7	87	16.5	112	17	Q8PS24 methanosaarc
8	86	16.3	952	11	Q8S444 mus musculu
9	85	16.2	301	4	Q8NTM5 homo sapien
10	84.5	16.1	108	16	Q8ZBR3 lactobacille
11	84.5	16.1	113	16	Q7V197 heliobacte
12	83.5	15.9	407	16	Q9A5Z6 caulobacter
13	83	15.8	108	16	Q8Y760 listeria mo
14	82.5	15.7	388	16	Q8FLC1 escherichia
15	82	15.6	113	17	Q8TKG9 mechanosarc
16	81.5	15.5	109	16	Q8CU94 staphylococc

17	81.5	15.5	182	16	Q81YS6 bacillus an
18	80.5	15.3	108	16	Q8BY20 lactobacill
19	80.5	15.3	192	16	Q9S5A1 rhizobium l
20	80.5	15.3	345	16	Q9ZUW5 rhizobium
21	80	15.2	589	16	Q9PEE7 xyloella fas
22	80	15.2	589	16	Q87EFO xyloella fas
23	79.5	15.1	956	16	Q7U948 synecococc
24	79	15.0	148	16	Q8XDT4 escherichia
25	79	15.0	314	17	Q9YXZ7 pyrococcus
26	79	15.0	394	16	Q8ZJZ6 salmoneilla
27	79	15.0	394	16	Q8ZJZ6 salmoneilla
28	78.5	14.9	388	16	Q8XA63 escherichia
29	78	14.8	443	16	Q98FMO rhizobium l
30	77.5	14.7	235	16	Q45Z35 bradyrhizob
31	77	14.6	388	16	Q83SR3 shigella fl
32	77	14.6	464	16	Q8DMD1 streptococc
33	76.5	14.5	348	16	Q8KDV2 chlorobium
34	76.5	14.5	440	16	Q8NMG2 corynebacte
35	76	14.4	102	16	Q50482 streptomyce
36	76	14.4	114	16	Q7MMJ7 bordetella
37	76	14.4	114	16	Q7WB31 bordetella
38	76	14.4	402	16	Q89RX2 bradyrhizob
39	76	14.4	459	16	Q88EE1 pseudomonas
40	76	14.4	543	3	Q86ZM8 podospira a
41	76	14.4	656	16	Q9KXK4 vibrio chol
42	75.5	14.4	107	16	Q8DVS8 streptococc
43	75.5	14.4	182	17	Q9YC04 aeropyrum p
44	75.5	14.4	506	13	Q8UVB7 brachydanio
45	75.5	14.4	506	13	Q7ZU13 brachydanio

ALIGNMENTS

Q8NTP7	PRELIMINARY;	PRT;	108 AA.
Q8NTP7	01-OCT-2002 (TREMURel. 22, Created)		
Q8NTP7	01-OCT-2002 (TREMURel. 22, Last sequence update)		
Q8NTP7	01-OCT-2003 (TREMURel. 25, Last annotation update)		
Q8NTP7	Hypothetical membrane protein Cg10259 (BrnE).		
Q8NTP7	Cg10259 OR BRNE.		
Q8NTP7	Corynebacterium glutamicum (Brevibacterium flavum).		
Q8NTP7	Corynebacterium glutamicum; Actinobacteridae; Actinomycetales;		
Q8NTP7	Corynebacteriaceae; Corynebacteriaceae; Corynebacterium.		
Q8NTP7	NCBI_TaxID=1718;		
Q8NTP7	SEQUENCE FROM N.A.		
Q8NTP7	STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;		
Q8NTP7	"Complete genomic sequence of Corynebacterium glutamicum ATCC 13032."		
Q8NTP7	Submitted (May-2002) to the EMBL/GenBank/DBJ databases.		
Q8NTP7	SEQUENCE FROM N.A.		
Q8NTP7	KEDINNE=22077265; PubMed=12081967;		
Q8NTP7	Kennethnecht N., Sahm H., Yen M.R., Patek M., Sailer M.H. Jr.,		
Q8NTP7	Reggeling L.,		
Q8NTP7	"Export of L-Isoleucine from Corynebacterium glutamicum: a Two-Gene-		
Q8NTP7	Encoded Member of a New Translocator Family."		
Q8NTP7	J. Bacteriol. 184:3947-3956(2002).		
Q8NTP7	EMBL; AP005274; BAB97652.1;		
Q8NTP7	EMBL; AF454053; AAM4685.1;		
Q8NTP7	InterPro: IPR008407; Asid.		
Q8NTP7	Pfam: PF05437; Asid. 1.		
Q8NTP7	Hypothetical protein; Complete proteome.		
Q8NTP7	SEQUENCE 108 AA; 11480 MW; 3A160B0842CC2CPD CRC64;		
Q8NTP7	Query Match	100.0%;	Score 526; DB 16; Length 108;
Q8NTP7	Best Local Similarity	100.0%;	Pred. No. 2.7e-45;
Q8NTP7	Matches 108; Conservative	0;	Mismatches 0; Indels 0; Gaps 0;
Q8NTP7	1 MTTDPSCLLVAVACAVITFALRAVPLILKPLRESQFVGGMAMMAGILAITASTFR	60	

```

FT DOMAIN 92 96 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 97 117 3 (POTENTIAL).
FT DOMAIN 118 127 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 128 148 4 (POTENTIAL).
FT DOMAIN 149 156 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 157 177 5 (POTENTIAL).
FT DOMAIN 178 182 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 183 203 6 (POTENTIAL).
FT DOMAIN 204 257 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 258 278 7 (POTENTIAL).
FT DOMAIN 279 293 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 294 314 8 (POTENTIAL).
FT DOMAIN 315 320 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 321 341 9 (POTENTIAL).
FT DOMAIN 342 367 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 368 388 10 (POTENTIAL).
FT DOMAIN 389 404 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 405 425 11 (POTENTIAL).
FT DOMAIN 426 438 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 439 459 12 (POTENTIAL).
FT DOMAIN 460 477 CYTOPLASMIC (POTENTIAL).
FT SITE 12 13 DILEUCINE INTERNALIZATION MOTIF (BY SIMILARITY).
FT CARBOHYD 350 350 N-LINKED (GLCNAC. .) (By similarity).
FT CONFLICT 39 39 S -> N (IN REF. 1).
FT CONFLICT 94 94 S -> A (IN REF. 2 AND 4).
FT CONFLICT 429 429 S -> N (IN REF. 1).
SQ SEQUENCE 477 AA, 51523 MW, A3753FB345455F9A CRC64;

Query Match 13.3%; Score 70; DB 1; Length 477;
Best Local Similarity 31.8%; Pred. No. 10;
Matches 27; Conservative 13; Mismatches 27; Indels 18; Gaps 4;

QY 29 ILKPLRESQVPGKMMMPAGILALITASTFRSNAL-----DLKTLFGILAVIT 79
DB 254 IKKPLIG--ISLMPQQLSGVNAIM---FYANSIFEBAKKDSLSAVTGIQVLFT 307
QY 80 VVAHLT---GGRRTLISVGATIVF 101
DB 308 AVAALIMDRAGRRLLALSLGVIMVF 332

RESULT 15
NUM PODAN STANDARD; PRT; 652 AA.
AC P20679;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 01-JUL-1993 (Rel. 26, Last annotation update)
DE NADH-ubiquinone oxidoreductase chain 5 (EC 1.6.5.3).
GN NDS.
OS Podospora anserina.
OG Mitochondrion.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariomycetidae; Sordariales; Lasiosphaeriaceae; Podospora.
OX NCBI_TaxID=5145;
RP SEQUENCE FROM N.A.
RC STRAIN=S, and A;
RX MEDLINE=90204555; Pubmed=2319602;
RA Cummings D.J., Michel F., Domenico J.M., McNally K.L.;
RT "DNA sequence analysis of the mitochondrial ND4L-ND5 gene complex
RT from Podospora anserina. Duplication of the ND4L gene within its
RT intron.";
RT J. Mol. Biol. 212:269-286(1990).
RN [2]
RP COMPLETE GENOME.
RC STRAIN=S;
RX MEDLINE=90291512; Pubmed=2357736;
RA Cummings D.J., McNally K.L., Domenico J.M., Matsuura E.T.;
RT "The complete DNA sequence of the mitochondrial genome of Podospora
RT anserina.";
RT Curr. Genet. 17:375-402(1990).

```

```

CC -1- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation.
CC The European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X55026; CA48798.1; -.
CC PIR; S09133; S09133.
CC InterPro; IPR003918; NADHub_oxred4.
CC InterPro; IPR003916; NADHub_oxred5.
CC InterPro; IPR001750; Oxidored_q1.
CC InterPro; IPR001516; Oxidored_q1_N.
CC Pfam; PF00661; Oxidored_q1; 1.
CC Pfam; PF00662; Oxidored_q1_N; 1.
CC PRINTS; PR01434; NADHOGNASE5.
CC PRINTS; PR01437; NUOXDRDTASE4.
CC Oxidoreductase; NAD; Ubiquinone; Mitochondrion.
CC SEQUENCE 652 AA; 72632 MW; 84CA887B75755F8 CRC64;

Query Match 13.3%; Score 70; DB 1; Length 652;
Best Local Similarity 25.9%; Pred. No. 13;
Matches 28; Conservative 19; Mismatches 37; Indels 24; Gaps 3;

QY 6 SCILVAVCAVTFMLRAVPFLILKPLRESQVPG-----KHAMMPAGILALITASTFR 60
DB 206 SKVLLIIGIC-----LLIGAMAKSSQVGLHWLPMVMBGPTPVSAIIHATWV 253
QY 61 SNADIKTLTFGLIAVAITVAHLGGRRTLISVGATIVFGLVNLV 108
DB 254 TAGVYLIMRSSPLIBYNSVTL-----LCWLIGAITTVFSSLIQIF 294

Search completed: April 16, 2004, 06:53:45
Job time : 7.01671 secs

```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL AF084229; AAD52041.1; -
 DR InterPro: IPR000276; GPCR_Rhodopsin.

DR Pfam: PF00001; 7tm.1; 1.

DR PRINTS: PR00237; G_PROTEIN_RECEP_F1_1; 1.

DR PROSITE: PS00237; G_PROTEIN_RECEP_F1_2; 1.
 DR G-protein coupled receptor; Transmembrane; Glycoprotein.

FT DOMAIN 1 32
 FT TRANSMEM 33 59
 FT DOMAIN 60 68
 FT TRANSMEM 69 89
 FT TRANSMEM 90 103
 FT TRANSMEM 104 125
 FT TRANSMEM 126 143
 FT TRANSMEM 144 164
 FT TRANSMEM 165 187
 FT TRANSMEM 188 215
 FT TRANSMEM 216 231
 FT TRANSMEM 232 259
 FT TRANSMEM 260 275
 FT TRANSMEM 276 293
 FT TRANSMEM 294 342
 FT DISULFID 102 180
 FT CARBOHYD 16 16

SO SEQUENCE 342 AA; 39273 MW; 5B5800379780662A CRC64; (POTENTIAL).

Query Match 13.3%; Score 70; DB 1; Length 342;
 Best Local Similarity 27.8%; Pred. No. 7.3;
 Matches 25; Conservative 12; Mismatches 37; Indels 16; Gaps 2;

QY 7 CILVVAVCAVITFAAY-----FLIKKPRSSQFVKMMNPAIGLAI 53
 DB 36 CMLVAVCGVGNLSLVVSTPHKQSLTDVFLVNLPLADLVFVCTLPFAVYAGIHNM 95
 QY 54 LRASTFSAIDKTLTF---GLIAVATV 80
 DB 96 VFGVWCKSLGTYITNFTSMILTCIV 125

RESULT 14
 GTR8 MOUSE STANDARD; PRT; 477 AA.
 AC Q9JIF3; Q9JUP4; Q9JUZ0;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Solute carrier family 2, facilitated glucose transporter, member 8
 GN (Glucose transporter type 8) (Glucose transporter type XI).
 DE SLC2A8 OR GLUT8 OR GLUTX1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 CC NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RP MEDLINE=20138191; PubMed=10671497; B;
 RA Ibberson M.R., Uldry M.A., Thorens B.;
 RT GLUTX1, a novel mammalian glucose transporter expressed in the
 RT central nervous system and insulin-sensitive tissues.";
 RL J. Biol. Chem. 275:4607-4612(2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RP TISSUE=Testis;
 RA MEDLINE=20283667; PubMed=10821868;
 RA Doerge H., Schnermann A., Bahrenberg C., Brauers A., Joest H.-G.;
 RT "GLUT8, a novel member of the sugar transport facilitator family with

RT glucose transport activity.";
 RL J. Biol. Chem. 275:16275-16280(2000).
 RN [3]
 RP SEQUENCE FROM N.A.
 RP STRAIN=129; TISSUE=Embryonic carcinoma;
 RC MEDLINE=20131902; PubMed=10960996;
 RA Carayannopoulos M.O., Chi M.M.-Y., Cui Y., Pingrethaus J.M.,
 RA McKnight R.A., Mueckler M., Devaskar S.U., Moley K.H.;
 RT GLUT8 is a glucose transporter responsible for insulin-stimulated
 RT glucose uptake in the placenta.
 RL Proc. Natl. Acad. Sci. U.S.A. 97:7313-7318(2000).
 RN [4]
 RP SEQUENCE FROM N.A.
 RP STRAIN=129/Ola; TISSUE=Spleen;
 RC MEDLINE=21547794; PubMed=11689004;
 RX Scheepers A., Doerge H., Joest H.-G., Schnermann A.;
 RT "Mouse GLUT8: genomic organization and regulation of expression in
 RT 3T3-L1 adipocytes by glucose."
 RL Biochem. Biophys. Res. Commun. 288:969-974(2001).
 CC -1- FUNCTION: Insulin-regulated facilitative glucose transporter.
 CC Binds cytochalasin B in a glucose-inhibitable manner. Seems to be
 CC a dual-specific sugar transporter as it is inhibitable by
 CC fructose.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Principally
 CC intracellular. May move between intracellular vesicles and the
 CC plasma membrane. The dileucine internalization motif is critical
 CC for intracellular sequestration (By similarity). Insulin induces a
 CC change in the intracellular localization and gives rise to
 CC insertion in the plasma membrane.
 CC -1- TISSUE SPECIFICITY: Highest level of expression in placenta and
 CC testis. Highly expressed in adult and pubertal testis, but not
 CC prepubertal testis. Lower levels of expression in brain, liver,
 CC heart, kidney, fat and skeletal muscle.
 CC -1- DEVELOPMENTAL STAGE: High expression in blastocysts.
 CC -1- INUNCTION: Inhibited under glucose deprivation.
 CC -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY. GLUCOSE
 CC TRANSPORTERS SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL AJ245936; CAB75719.1; -
 DR EMBL Y17802; CAB89815.1; -
 DR EMBL AF232061; AAF78366.1; -
 DR EMBL AJ413951; CAC98690.1; -
 DR WGD; MG1:1860103; SLC2A8.
 DR GO; GO:0005887; C:integral to plasma membrane; IDA.
 DR GO; GO:0005336; F:glucose binding; IDA.
 DR GO; GO:0005355; F:glucose transporter activity; IDA.
 DR GO; GO:0015758; P:glucose transport; IDA.
 DR GO; GO:0008286; P:insulin receptor signaling pathway; IDA.
 DR GO; GO:0001665; P:response to hypoxia; IDA.
 DR InterPro: IPR007114; WFS.
 DR InterPro: IPR005828; Sub_transporter.
 DR InterPro: IPR005829; Sug_transporter.
 DR InterPro: IPR003663; Sugar_transp.
 DR Pfam: PF00083; sugar tr; 1.
 DR PRINTS: PR00171; SUGRTRANSPORT.
 DR TIGRPFAM: TIGR00679; Sp. 1.
 DR PROSITE: PS00580; MFS; 1.
 DR PROSITE: PS00216; SUGAR_TRANSPORT_1; 1.
 DR PROSITE: PS00217; SUGAR_TRANSPORT_2; 1.
 KM Transpote; Sugar transport; Transmembrane; Glycoprotein;
 KM Multigene family.
 FT DOMAIN 1 26
 FT TRANSMEM 26 46
 FT TRANSMEM 47 70
 FT TRANSMEM 71 91

RT immunodeficiency viruses.";
 RL Nature 388:296-300(1997).
 RN [2]
 RP SEQUENCE FROM N.A., AND VARIANT ALA-25.
 RX MEDLINE=97311099; PubMed=9166430;
 RA Liao P., Alkhatib G., Peden K.W.C., Sharma G., Berger E.A.,
 RA Farber J.M.,
 RA "STRL3, a novel chemokine receptor-like protein, functions as a
 RT fusion cofactor for both macrophage-tropic and T cell line-tropic
 RT HIV-1.";
 RL J. Exp. Med. 185:2015-2023(1997).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX TISSUE=Blood;
 RX MEDLINE=97431687; PubMed=9285716;
 RA Loetscher M., Amara A., Oberlin E., Brass N., Legler D.F.,
 RA Loetscher P., D'Apuzzo M., Meese E.U., Rouselet D., Virelizier J.L.,
 RA Baggiolini M., Arenzana-Seisdedos F., Moser B.,
 RA "XCR4, a putative chemokine receptor selectively expressed in
 RT activated T cells, exhibits HIV-1 coreceptor function.";
 RL Curr. Biol. 7:652-660(1997).
 RN [4]
 RP SEQUENCE FROM N.A.
 RX Kopatz S.A., Aronstam R.S., Sharma S.V.,
 RT "cDNA clones of human proteins involved in signal transduction
 RT sequenced by the Guthrie cDNA resource center (www.cdna.org).";
 RL Submitted (JUN-2003) to the EMBL/Genbank/DBJ databases.
 RN [5]
 RP SEQUENCE FROM N.A.
 RX TISSUE=Pancreas;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Klausner R.D., Collins F.S., Wagner L., Shemen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Donald W.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ustin T.B., Toshlyuk S., Carninci P., Prange C.,
 RA Raha S.S., Loggiano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny K.C., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Butterfield Y.S.N., Krzywnicki M.I., Skalska U., Smalhus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 CC -1- FUNCTION: Receptor for the C-X-C chemokine CXCL16. Used as a
 CC coreceptor by SIVs and by strains of HIV-2 and m-tropic HIV-1.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- TISSUE SPECIFICITY: Expressed in lymphoid tissues and activated T
 CC cells.
 CC -1- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.sib.ch/announce/
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: AF007545; AAB64221.1; -
 CC EMBL: U73529; AAB61456.1; -
 CC EMBL: U73531; AAB61457.1; -
 CC EMBL: Y13248; CAA73698.1; -
 CC EMBL: AY322543; AAB84356.1; -
 CC EMBL: BC033584; AAB33584.1; -
 CC Genew; HGNC:16647; CXCR6.

DR MM; 605163; -
 DR GO; GO:0005887; C: integral to plasma membrane; TAS.
 DR GO; GO:0015026; F: coreceptor activity; TAS.
 DR GO; GO:0004930; F: G-protein coupled receptor activity; TAS.
 DR GO; GO:0007186; F: G-protein coupled receptor protein signalin...; TAS.
 DR GO; GO:0008166; P: viral replication; TAS.
 DR InterPro; IPR000276; GPCR_Rhodopsn.
 DR Pfam; PF00001; 7tm.1; 1.
 DR PRINTS; PR00237; GPCRHOPOSPN.
 DR PROSITE; PS00252; G-PROTEIN_RECEP_F1_1; 1.
 DR PROSITE; PS00252; G-PROTEIN_RECEP_F1_2; 1.
 KW G-protein coupled receptor; Transmembrane; Glycoprotein; Polymorphism.
 FT DOMAIN 1
 FT TRANSMEM 33 59
 FT DOMAIN 60 68
 FT TRANSMEM 69 89
 FT DOMAIN 90 103
 FT TRANSMEM 104 125
 FT DOMAIN 126 143
 FT TRANSMEM 144 164
 FT TRANSMEM 165 187
 FT TRANSMEM 188 215
 FT DOMAIN 216 231
 FT TRANSMEM 232 259
 FT DOMAIN 260 275
 FT TRANSMEM 276 293
 FT DOMAIN 294 342
 FT DISULFID 102 180
 FT CARBOHYD 16 16
 FT VARIANT 25 25
 FT
 SQ SEQUENCE 342 AA; 39280 MM; 9FBC02555601082E CRC64;
 Query Match 13.3%; Score 70; DB 1; Length 342;
 Best Local Similarity 27.8%; Pred. No. 7.3;
 Matches 25; Conservative 12; Mismatches 37; Indels 16; Gaps 2;
 CQ 7 CILVAVCAVITFAKAV-----FLIKPAREQFYKXMMNPAGILAI 53
 DB 36 CMYLVAVFCVGVNSLVLVISTYHQLQSLDVFVNLPLADLVFCTLPFMYAVAGIHEW 95
 CQ 54 LTASTFRSNAIDKLTTF---GLIYAVITV 80
 DB 96 VFGWCKSLIGIYITINFTSMILTCITV 125
 RESULT 13
 AC CCR6_PANTR STANDARD; PRT; 342 AA.
 AC Q9TVI6;
 DT 15-MAR-2004 (Rel. 43, Created)
 DT 15-MAR-2004 (Rel. 43, Last sequence update)
 DE 15-MAR-2004 (Rel. 43, Last annotation update)
 DE C-X-C Chemokine receptor type 6 (CXCR-6) (CXCR-6) (G protein-coupled
 DE receptor bonzo) (G protein-coupled receptor STRL33).
 GN CXCR6 OR BONZO OR STRL33.
 OS Pan troglodytes (Chimpanzee).
 OC Eukaryota; Metazoa; Chordata;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pan.
 OK NCBI_Taxid=9598;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99433499; PubMed=10505680;
 RA Brusee A., Pretet J.-L., Girard M., Butor C.,
 RA "Sequences and predicted structures of chimpanzee STRL33 (Bonzo) and
 RT gp15 (BOB).";
 RT AIDS Res. Hum. Retroviruses 15:1315-1319(1999).
 CC -1- FUNCTION: Receptor for the C-X-C chemokine CXCL16. Used as a
 CC coreceptor by SIVs and by strains of HIV-2 and m-tropic HIV-1.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration

DR PIR, S04254; G31277.
 DR InterPro: IPR007114; MFS.
 DR InterPro: IPR005828; Sub_transporter.
 DR InterPro: IPR005829; Sug_transporter.
 DR InterPro: IPR003663; Sugar_transp.
 DR Pfam: PF00083; sugar tr; 1.
 DR PRINTS: PR00171; SUGTRNSPORT.
 DR TIGRfam: TIGR00879; Sp; 1.
 DR PROSITE: PS00850; MFS; 1.
 DR PROSITE: PS00216; SUGAR_TRANSPORT_1; 1.
 DR PROSITE: PS00217; SUGAR_TRANSPORT_2; 1.
 DR Transmembrane; Transport; Quinate metabolism; Glycoprotein.
 FT DOMAIN 1 26
 FT TRANSSEM 27 47
 FT TRANSSEM 48 74
 FT TRANSSEM 75 95
 FT TRANSSEM 96 98
 FT TRANSSEM 99 119
 FT TRANSSEM 120 131
 FT TRANSSEM 132 152
 FT TRANSSEM 153 160
 FT TRANSSEM 161 181
 FT TRANSSEM 182 195
 FT TRANSSEM 196 216
 FT TRANSSEM 217 285
 FT TRANSSEM 286 306
 FT TRANSSEM 307 327
 FT TRANSSEM 328 349
 FT TRANSSEM 350 352
 FT TRANSSEM 353 373
 FT TRANSSEM 374 389
 FT TRANSSEM 390 410
 FT TRANSSEM 411 435
 FT TRANSSEM 436 456
 FT TRANSSEM 457 458
 FT TRANSSEM 459 479
 FT TRANSSEM 480 537
 FT CARBOHYD 184 194
 FT SEQUENCE 537 AA; 60103 MW; 9AC63400FC164F3 CRC64;
 Query Match 13.7%; Score 72; DB 1; Length 537;
 Best Local Similarity 35.4%; Pred. No. 7.1;
 Matches 35; Conservative 16; Mismatches 26; Indels 22; Gaps 9;

QY 26 PFLIKPRLRSQ---FVGKKA-MWM-PAGILAI--LTASFRSNAID-----LKTLPFG 72
 DB 273 PFLSIKQ-RKQWRFIFGMLFPGWNGSGINAINVSPYFRSIGITGDTGFLTIGIFG 331
 QY 73 LIAVAIVTVVAHL---LGRRTLLSVCAG--GTV--FVG 103
 DB 332 VVKVLTIIWLMVLVDVGRRIILFIGAAGSLQWFFIG 370

RESULT 11

CCR6 CERAE STANDARD; PRT; 342 AA.
 ID 016983;
 AC 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE C-X-C chemokine receptor type 6 (CXCR-6) (G protein-coupled receptor bonzo).
 DE C-X-C chemokine receptor type 6 (CXCR-6) (G protein-coupled receptor bonzo).
 GN CCR6 OR BONZO.
 OS Cercopithecus aethiops (Green monkey) (Primate).
 OC Eukaryota; Metazoa; Chordata; Ctenista; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 OC Cercopithecinae; Cercopithecus.
 OX NCBI_TaxID=9534;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97373958; PubMed=9230441;
 RA Deng H.K., Unutmaz D., Kewalramani V.N., Littman D.R.;
 RT "Expression cloning of new receptors used by simian and human

RT Immunodeficiency viruses."
 RL Nature 388:296-300(1997).
 CC -1- FUNCTION: Receptor for the C-X-C chemokine CXCL16. Used as a
 CC coreceptor by SIVs and by strains of HIV-2 and m-tropic HIV-1.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb.ch/announce/>
 CC or send an email to license@isb.ch).
 CC
 DR EMBL, AF007859; AAB64225.1; -
 DR InterPro: IPR000276; GPCR_Rhodopsn.
 DR Pfam: PF00001; 7tm_1; 1.
 DR PRINTS: PR00237; GPCRHDOPSN.
 DR PROSITE: PS00237; G_PROTEIN_RECEP_F1_1; 1.
 DR PROSITE: PS02622; G_PROTEIN_RECEP_F1_2; 1.
 DR G-protein coupled receptor; Transmembrane; Glycoprotein.
 FT DOMAIN 1 32
 FT TRANSSEM 33 39
 FT TRANSSEM 60 68
 FT TRANSSEM 69 89
 FT TRANSSEM 90 103
 FT TRANSSEM 104 125
 FT TRANSSEM 126 143
 FT TRANSSEM 144 164
 FT TRANSSEM 165 187
 FT TRANSSEM 188 215
 FT TRANSSEM 216 231
 FT TRANSSEM 232 259
 FT TRANSSEM 260 275
 FT TRANSSEM 276 293
 FT TRANSSEM 294 342
 FT TRANSSEM 342 342
 FT DISULFID 102 180
 FT CARBOHYD 16 16
 FT SEQUENCE 342 AA; 39226 MW; 6CBFB389CE5919E CRC64;
 Query Match 13.3%; Score 70; DB 1; Length 342;
 Best Local Similarity 27.8%; Pred. No. 7.3;
 Matches 25; Conservative 11; Mismatches 38; Indels 16; Gaps 2;

RESULT 12

CCR6 HUMAN STANDARD; PRT; 342 AA.
 ID 000574; 000575;
 AC 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE C-X-C chemokine receptor type 6 (CXCR-6) (G protein-coupled receptor bonzo) (G protein-coupled receptor STR333).
 DE C-X-C chemokine receptor type 6 (CXCR-6) (G protein-coupled receptor bonzo) (G protein-coupled receptor STR333).
 GN CCR6 OR BONZO OR STR333 OR TYMOTR.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Ctenista; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97373958; PubMed=9230441;
 RA Deng H.K., Unutmaz D., Kewalramani V.N., Littman D.R.;
 RT "Expression cloning of new receptors used by simian and human

```
FT DOMAIN 497 538 CYTOPLASMIC (BY SIMILARITY).
FT TRANSMEM 539 557 7 (BY SIMILARITY).
FT DOMAIN 558 572 PERIPLASMIC (BY SIMILARITY).
FT TRANSMEM 873 892 8 (BY SIMILARITY).
FT DOMAIN 893 898 CYTOPLASMIC (BY SIMILARITY).
FT TRANSMEM 899 918 9 (BY SIMILARITY).
FT DOMAIN 919 924 PERIPLASMIC (BY SIMILARITY).
FT TRANSMEM 925 946 10 (BY SIMILARITY).
FT DOMAIN 947 974 CYTOPLASMIC (BY SIMILARITY).
FT TRANSMEM 975 993 11 (BY SIMILARITY).
FT DOMAIN 994 1006 PERIPLASMIC (BY SIMILARITY).
FT TRANSMEM 1007 1029 12 (BY SIMILARITY).
FT DOMAIN 1030 1049 CYTOPLASMIC (BY SIMILARITY).
SQ SEQUENCE 1049 AA; 113573 MW; 19670E3C4CC29055 CRC64;

Query Match 14.0%; Score 73.5; DB 1; Length 1049;
Best Local Similarity 19.4%; Pred. No. 9.5;
Matches 26; Conservative 22; Mismatches 33; Indels 53; Gaps 3;

QY 5 FSCILLVAVCAVITFALRAVPELILKPLRSQVGRAMMPGIALILASTFR--- 60
DB 877 VAISLIVFLCLALYESSMSIPFSVM-----LVVPLGVIGALLAATFRGLTN 923
QY 61 -----SNAI-----DKLTTFGLIAVATVVAHL 84
DB 924 DVYFVGVGLITTTGLSAKNILIVERPAKDMKEKGLIEATLIDAVRMLRPLMTSLAFI 983
QY 85 LGGRTLLSVAGT 98
DB 984 LGVMPLELVISTAGS 997

RESULT 9
Y420_NEIMA STANDARD; PRT; 227 AA.
ID Y420_NEIMA
AC Q9JRI0;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
GN Hypothetical protein NMA0420/NMB2020.
OS Neisseria meningitidis (serogroup A), and
OS Neisseria meningitidis (serogroup B).
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC Neisseriaceae; Neisseria.
OX NCBI_TaxID=65699, 491;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=22491 / Serogroup A / Serotype 4A;
RX MEDLINE=20222556; PubMed=10761919;
RA Parhill J., Achtman M., James K.D., Bentley S.D., Churcher C.,
RA Klee S.R., Morelli G., Basham D., Brown D., Chillingworth T.,
RA Davies R.M., Davis P., Devlin K., Feltwell T., Hamlin N., Holtroyd S.,
RA Jagers K., Leather S., Moule S., Mungall K., Quail M.A.,
RA Rajadaram M.A., Rutherford K.M., Simmonds M., Skelton J.,
RA Whitehead S., Spratt B.G., Barrett B.G.;
RA "Complete DNA sequence of a serogroup A strain of Neisseria
RT meningitidis 22491."
RT Nature 404:502-506(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=MC58 / Serogroup B;
RX MEDLINE=20175755; PubMed=10710307;
RA Tetteilin H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E.,
RA Bisen J.A., Kechum K.A., Hood D.W., Peden J.F., Dodson R.J.,
RA Nelson W.C., Gilm M.L., Deboy R., Peterson J.D., Hickey E.K.,
RA Heft D.H., Salberg S.L., White O., Fleischmann R.D., Dougherty B.A.,
RA Mason T., Ciecko A., Parksey D.S., Blair B., Citterone H., Clark E.B.,
RA Colton M.D., Urdachack T.R., Khouri H., Qin H., Vamathevan J.,
RA Gill J., Scarlato V., Maestroni V., Piazza M., Grandi G., Sun J.,
RA Smith H.O., Fraser C.M., Moxon E.R., Rappaport J., Venter J.C.;
RT "Complete genome sequence of Neisseria meningitidis serogroup B strain
RT MC58.";
```

```
RL Science 287:1809-1815(2000).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AL162753; CAB83719.1;
DR EMBL; AE002552; AAF42343.1;
DR PIR; G81014; G81014.
DR TIGR; NMB2020;
DR InterPro; IPR006214; UPF0005.
DR Pfam; PF01027; UPF0005; 1.
KW Hypothetical protein; Transmembrane; Complete proteome.
FT TRANSMEM 25 45 POTENTIAL.
FT TRANSMEM 49 69 POTENTIAL.
FT TRANSMEM 80 100 POTENTIAL.
FT TRANSMEM 111 131 POTENTIAL.
FT TRANSMEM 144 164 POTENTIAL.
FT TRANSMEM 165 185 POTENTIAL.
FT TRANSMEM 201 221 POTENTIAL.
SQ SEQUENCE 227 AA; 24454 MW; 0E24A04F9676CDBD CRC64;

Query Match 13.7%; Score 72; DB 1; Length 227;
Best Local Similarity 31.9%; Pred. No. 3.2;
Matches 22; Conservative 14; Mismatches 21; Indels 12; Gaps 3;

QY 49 GILAILTASTF-----RSNAIDL---KTLTFGLIAVATVVAHL-LGRRITLSVGA 96
DB 114 GIAAMTPRAVFLTMSALARRTRIDNNALGRFLTVGAVIIMVAVANLFGIPALALTISA 173
QY 97 GTIVFVGLV 105
DB 174 GFVLFSSLM 182

RESULT 10
QAY_NEUCR STANDARD; PRT; 537 AA.
ID QAY_NEUCR
AC P11636;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Quinate permease (Quinate transporter).
GN QAY-.
OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
OX NCBI_TaxID=5141;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=74-OR23-1A / FGSC 987;
RX MEDLINE=89293848; PubMed=2525625;
RA Geever R.F., Hulet L., Baum J.A., Tyler B.M., Patel V.B.,
RA Rutledge B.J., Case M.B., Giles N.H.;
RA "DNA sequence, organization and regulation of the qa gene cluster of
RT Neurospora crassa."
RT J. Mol. Biol. 207:15-34(1989).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: Belongs to the sugar transporter family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X14603; CAA32752.1;
DR
```


OC Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 OC NCBI_Taxid=9913;
 RN [1]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RC TISSUE=Brain;
 RX MEDLINE=9298382; PubMed=2472670;
 RA Krupinski J., Coussen F., Bakalyar H.A., Tang W.-J.,
 RA Feinstein P.G., Orth K., Slaughter C., Reed R.R., Gilman A.G.,
 RA "adenyl cyclase amino acid sequence: possible channel- or
 RT transporter-like structure."
 RT Science 244:1558-1564(1989).
 RL [2]
 RL 3D-STRUCTURE MODELING OF 295-450; 861-936 AND 950-1045.
 RP MEDLINE=98054247; PubMed=9391039;
 RA Liu Y., Ruoho A.E., Rao V.D., Hurley J.H.,
 RA "Catalytic mechanism of the adenylyl and guanylyl cyclases: modeling
 RT and mutational analysis."
 RT Proc. Natl. Acad. Sci. U.S.A. 94:13414-13419(1997).
 CC -1- FUNCTION: This is a calmodulin-sensitive adenylyl cyclase. May be
 CC involved in regulatory processes in the central nervous system. It
 CC may play a role in memory acquisition and learning.
 CC -1- CATALYTIC ACTIVITY: ATP = 3',5'-cyclic AMP + diphosphate.
 CC -1- COFACTOR: Binds 2 magnesium ions per subunit (By similarity).
 CC -1- ENZYME REGULATION: Activated by calcium/calmodulin. Inhibited by
 CC the G protein beta and gamma subunit complex.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- TISSUE SPECIFICITY: Brain.
 CC -1- DOMAIN: Composed of two homologous domains.
 CC -1- SIMILARITY: Belongs to the adenylyl cyclase class-4/guanylyl
 CC cyclase family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: M25579; AAA79957.1; -
 CC PIR: A41350; A41350.
 CC PDB: 1AMK; 28-JAN-98.
 CC InterPro: IPR001054; G_cyclase.
 CC Pfam: PFO0211; guanylate_cyc; 2.
 CC SMART: SMO0044; CYCG; 2.
 CC PROSITE: PS00452; GUANYLATE_CYCLASES_1; 2.
 CC PROSITE: PS50125; GUANYLATE_CYCLASES_2; 2.
 CC Lyase; CAMP biosynthesis; Transmembrane; Glycoprotein; Repeat;
 CC Calmodulin-binding; Metal-binding; Magnesium; 3D-structure.
 KW DOMAIN 1
 FT FT 1 65
 FT FT 66 86
 FT FT 87 110
 FT FT 111 147
 FT FT 148 174
 FT FT 175 204
 FT FT 205 236
 FT FT 237 612
 FT FT 613 633
 FT FT 634 657
 FT FT 658 696
 FT FT 697 726
 FT FT 727 747
 FT FT 748 775
 FT FT 776 797
 FT FT 798 1134
 FT FT 1135 522
 FT FT 523 1050
 FT FT 1051 34
 FT FT 35 310
 FT FT 311 311
 FT METAL
 FT METAL 311 311

FT METAL 354 354
 FT CARBOHYD 706 706 MAGNESIUM 1 AND 2 (BY SIMILARITY);
 FT STRAND 298 300 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT STRAND 304 310
 FT HELIX 312 315
 FT TURN 316 317
 FT TURN 320 321
 FT HELIX 324 337
 FT TURN 338 339
 FT HELIX 340 342
 FT TURN 344 345
 FT STRAND 348 348
 FT STRAND 355 360
 FT TURN 367 367
 FT HELIX 368 391
 FT TURN 392 392
 FT STRAND 398 404
 FT STRAND 407 410
 FT STRAND 418 419
 FT STRAND 423 423
 FT HELIX 424 434
 FT TURN 434 434
 FT HELIX 444 447
 FT TURN 448 449
 FT STRAND 863 864
 FT STRAND 870 873
 FT HELIX 879 883
 FT HELIX 888 890
 FT TURN 891 893
 FT HELIX 894 912
 FT TURN 914 915
 FT STRAND 924 925
 FT TURN 926 927
 FT STRAND 928 929
 FT STRAND 931 933
 FT TURN 951 952
 FT HELIX 953 971
 FT TURN 972 973
 FT STRAND 983 985
 FT STRAND 990 992
 FT STRAND 998 1000
 FT HELIX 1004 1014
 FT TURN 1015 1015
 FT HELIX 1025 1030
 FT HELIX 1031 1033
 FT TURN 1034 1035
 FT SO SEQUENCE 1134 AA; 123978 MW; CCA410BCE224DF3 CRC64;
 SO 1134 AA; 123978 MW; CCA410BCE224DF3 CRC64;
 Query Match 14.3%; Score 75; DB 1; Length 1134;
 Best Local Similarity 28.6%; Pred. No. 7.2;
 Matches 30; Conservative 20; Mismatches 45; Indels 10; Gaps 5;
 QY 12 VAVCAVTFPALRAVPFPIILCPLEBSQVGVGMAMVPAIGILAITASTFRESNAI-----DLK 67
 DB 127 LALFSTFALLCFPALGPAGAHGAAVPAITADGVQMLLV-TFVSALLPVRSL 165
 QY 68 TLTFGLIAVA--ITVVAHLGGRITLL--SVGAGTIVFVGLVNLF 108
 DB 186 AIGGIVVAASHLVLTATLVPAKRPRLMTLGAVALFLG-VVNY 229
 RESULT 7
 TRPP_CLOAB STANDARD; PRT; 171 AA.
 AC Q97D63;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Probable tryptophan transport protein.
 GN TRPP OR CAC3617.
 OS Clostridium acetobutylicum.
 OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
 OC Clostridium.

RA Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,
 RA Raebio D., Buckles E.L., Ikon S.-R., Boutin A., Hackett J., Stroud D.,
 RA Mobley G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,
 RA Mobley H.L.T., Donnenberg M.S., Blattner F.R.;
 RT "Extensive mosaic structure revealed by the complete genome sequence
 RT of uropathogenic *Escherichia coli*.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024(2002).
 [4]
 RN SEQUENCE FROM N.A.
 RC SPECIES=S.flexneri; STRAIN=301 / Serotype 2a;
 RX MEDLINE=22272406; PubMed=12384590;
 RA Jin Q., Yuan Z., Xu U., Wang Y., Shen Y., Lu W., Wang J., Liu H.,
 RA Yang U., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,
 RA Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,
 RA Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,
 RA Yu J.;
 RT "Genome sequence of *Shigella flexneri* 2a: insights into pathogenicity
 RT through comparison with genomes of *Escherichia coli* K12 and O157.";
 RL Nucleic Acids Res. 30:4432-4441(2002).
 [5]
 RN SEQUENCE FROM N.A.
 RC SPECIES=S.flexneri; STRAIN=2457T / ATCC 700930 / Serotype 2a;
 RX MEDLINE=22590274; PubMed=12704152;
 RA Wei J., Goldberg M.B., Burland V., Venkatesan M.M., Deng W.,
 RA Fournier G., Mayhew G.F., Plunkett G. III, Rose D.J., Darling A.,
 RA Mau B., Perna N.T., Payne S.M., Runyen-Janecky L.J., Zhou S.,
 RA Schwartz D.C., Blattner F.R.;
 RT "Complete genome sequence and comparative genomics of *Shigella*
 RT flexneri serotype 2a strain 2457T.";
 RL Infect. Immun. 71:2775-2786(2003).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -1- SIMILARITY: STRONG, TO H.INFLUENZA H1074.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AEO00274; AAC74859.1; -;
 DR EMBL; D90822; BAA15588.1; -;
 DR EMBL; D90823; BAA15590.1; -;
 DR EMBL; AE016761; AAN80653.1; -;
 DR EMBL; AE015167; AAN43034.1; -;
 DR EMBL; AE016983; AAP16929.1; -;
 DR PIR; E64939; E64939;
 DR EcoGene; EG13496; Yeal.
 DR InterPro; IPR007382; DUF441.
 DR Pfam; PF04284; DUF441.1.
 KM Hypothetical protein; Transmembrane; Complete proteome.
 FT TRANSMEM 14 34 POTENTIAL.
 FT TRANSMEM 51 71 POTENTIAL.
 FT TRANSMEM 86 106 POTENTIAL.
 FT TRANSMEM 121 141 POTENTIAL.
 SQ SEQUENCE 148 AA; 15255 MW; C24E91E1BA51FB93 CRC64;

Query Match 15.0%; Score 79; DB 1; Length 148;
 Best Local Similarity 22.5%; Pred. No. 0.43;
 Matches 27; Conservative 20; Mismatches 49; Indels 24; Gaps 2;

ID YH37_HAEIN STANDARD; PRT; 109 AA.
 AC P44301;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Hypothetical protein H11737.
 GN H11737.
 OS Haemophilus influenzae.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
 OC Pasteurellaceae; Haemophilus.
 CX NCBI_TaxID=727;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Rd / KM20 / ATCC 51907;
 RX MEDLINE=95350630; PubMed=7542800;
 RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
 RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
 RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
 RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
 RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
 RA Uterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
 RA Fine L.D., Frichman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
 RA Gann C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
 RA Venter J.C.;
 RT "Whole-genome random sequencing and assembly of *Haemophilus influenzae*
 RT Rd.";
 RL Science 269:496-512(1995).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -1- SIMILARITY: BELONGS TO THE AZLD/H11737/HP1330 FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; U32846; AAC23381.1; -;
 DR PIR; E64041; E64041.
 DR TIGR; H11737; -;
 DR InterPro; IPR008407; AZLD.
 DR Pfam; PF05437; AZLD.1.
 KM Hypothetical protein; Transmembrane; Complete proteome.
 FT TRANSMEM 7 27 POTENTIAL.
 FT TRANSMEM 37 57 POTENTIAL.
 FT TRANSMEM 63 83 POTENTIAL.
 SQ SEQUENCE 109 AA; 12423 MW; BCC25DB6F95005DF CRC64;

Query Match 14.4%; Score 75.5; DB 1; Length 109;
 Best Local Similarity 20.2%; Pred. No. 0.71;
 Matches 21; Conservative 27; Mismatches 47; Indels 9; Gaps 4;

12 VAVCAVTFPALRAVPFLIL--KPLRESQFVGKAMMMPAGILAILASTFRSNAID 71
 24 VAVSILVITVRVPLSTFPFWIEKQSLGIIITLTVNAPLASGLPSTLHSFLNW 83
 72 -GLIAVAITVAHLIGRRLL-----SYGACTIVFVGLVNI 107
 84 KSLVAIVAVIVSLGGRVTLMSQPOLVAGLVGVTVLGVAFRGVGPFLIAGHVS 143

9 LTVAVCAVTFPALRAVPFLIL--KPLRESQFVGKAMMMPAGILAILASTFRSNAID 65
 7 IITIGICIVAVQFRLRPFVFPVNRPI--PQYIRYIGKVLFPAMFGLVVYCYKN--IE 62
 66 LKTLTFLG--IAVAITVAHLIGRRLLSYGACTIVFVGLVNI 107
 63 ILTVGIGIPDLIAGIVLGLHFVKKNFSLAVGTLTFMALVOL 106

RESULT 6
 CVAL BOVIN STANDARD; PRT; 1134 AA.
 AC P19754;
 DT 01-FEB-1991 (Rel. 17, Created)
 DT 01-FEB-1991 (Rel. 17, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Adenylate cyclase, type I (EC 4.6.1.1) (ATP pyrophosphate-lyase)
 DE (Ca(2+)/calmodulin activated adenyl1 cyclase).
 GN ADCYL.
 OS Bos taurus (Bovine).

FT SITE 225 225 IMPORTANT FOR PH SENSOR.
 SQ SEQUENCE 368 AA; 41355 MM; B508BD25EE9130 CRC64;
 Query Match 15.5%; Score 81.5; DB 1; Length 388;
 Best Local Similarity 30.3%; Pred. No. 0.6;
 Matches 27; Conservative 18; Mismatches 35; Indels 9; Gaps 3;
 DB 134 IAFALGVALTGSFVPLKTLF--MALATIDGATITITLFTNDLSASLGAVALI 191
 18 ITPALRAAPFLIK-PLAESQFVGKMMMPAGIATITASTFRSMIDKTLFGLIAY 76
 77 AITVVAHLGGRTLLSVAGTIVEGLV 105
 192 AVLAVALNCGART-----GVYILVGV 214
 RESULT 3
 YE1F SCHPO STANDARD; PRT; 628 AA.
 ID YE1F SCHPO STANDARD; PRT; 628 AA.
 AC 013879;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Putative transporter CIB3.15C.
 GN SPAC1B3.15C.
 OS Schizosaccharomyces pombe (fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomycetes.
 NC NCB1_taxid=4896;
 RX MEDLINE=21848401; PubMed=11859360;
 RC STRAIN=972;
 RA Wood V., Gilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
 RA Sgouras J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
 RA Collins M., Connor R., Cronin A., Davis P., Felwell T., Frazer A.,
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
 RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jazgale K.,
 RA James K., Jones L., Jones M., Leach S., McDonald S., McLean J.,
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitch E.,
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
 RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
 RA Woodward J., Volkart G., Aert R., Robben J., Grynopre B.,
 RA Wellens I., Vancraets E., Rieger M., Schaefer M., Mueller-Huer S.,
 RA Gabel C., Fuchs M., Fritz C., Holzer H., Moestl D., Hilbert H.,
 RA Borzym K., Langer I., Beck A., Lehnach H., Reinhardt R., Pohl T.M.,
 RA Eger P., Zimmermann W., Medler H., Wandt R., Purnelle B.,
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Delaure V., Nottier S.,
 RA Gallbert F., Aves S.J., Xiang Z., Hunt C., Moore K., Huret S.M.,
 RA Lucas M., Rochet M., Gallard C., Tallada V.A., Garçon A., Thode G.,
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forzhang S.L.,
 RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
 RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
 RA "The genome sequence of Schizosaccharomyces pombe".
 RT Nature 415:871-880(2002).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -1- SIMILARITY: Belongs to the allantoate permease family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC EMBL: Z98598; CAB1241.1; -
 DR EIR; T38034; T38034.

DR GenedB Spombe; SPAC1B3.15C; -
 KW Hypothetical protein; Transmembrane; Transport.
 FT TRANSMEM 157 177 POTENTIAL.
 FT TRANSMEM 203 223 POTENTIAL.
 FT TRANSMEM 230 250 POTENTIAL.
 FT TRANSMEM 259 279 POTENTIAL.
 FT TRANSMEM 294 314 POTENTIAL.
 FT TRANSMEM 324 344 POTENTIAL.
 FT TRANSMEM 417 437 POTENTIAL.
 FT TRANSMEM 454 474 POTENTIAL.
 FT TRANSMEM 483 503 POTENTIAL.
 FT TRANSMEM 511 531 POTENTIAL.
 FT TRANSMEM 542 562 POTENTIAL.
 FT TRANSMEM 583 603 POTENTIAL.
 FT DOMAIN 92 102 POLY-SER.
 FT DOMAIN 385 388 POLY-GLU.
 SQ SEQUENCE 628 AA; 69076 MM; E1040C37445A6D36 CRC64;
 Query Match 15.2%; Score 80; DB 1; Length 628;
 Best Local Similarity 29.2%; Pred. No. 1.3;
 Matches 33; Conservative 22; Mismatches 46; Indels 12; Gaps 5;
 DB 448 NFSVTVSLVAPWIF--DAIALIVLPHD-RFHKRLTFVSGCLFVLAGLITTFVS 504
 62 NA-----IDKTLTFGL--TAAITVVAHLGGRTLLSVAGTIVEGLVNL 107
 505 NVMGKRVYGLIILFGIGPTVPIIMTWSSAMGSHGDVGAAGLAIVSGIGNL 557
 RESULT 4
 YEAL ECOLI STANDARD; PRT; 148 AA.
 ID YEAL ECOLI STANDARD; PRT; 148 AA.
 AC P76240; O07965; O07967;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Hypothetical protein yeast.
 GN YEAL OR B1789 OR C2194 OR SF7435 OR S1550.
 OS Escherichia coli.
 OS Escherichia coli O6, and
 OS Shigella flexneri.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.
 NC NCB1_taxid=562, 217992, 623;
 RX MEDLINE=97251358; PubMed=8097040;
 RC SPECIES=E.coli; STRAIN=KL2 / MG1655;
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Shao Y.;
 RA "The complete genome sequence of Escherichia coli K-12".
 RT Science 277:1453-1474(1997).
 CC [2]
 CC SEQUENCE FROM N.A.
 CC SPECIES=E.coli; STRAIN=KL2;
 CC MEDLINE=97251358; PubMed=8097040;
 CC Itoh T., Aliba H., Baba T., Fujita K., Hayashi K., Inada T., Isono K.,
 CC Kasai H., Kimura S., Kitakawa M., Kitagawa M., Makino K., Miki T.,
 CC Mizobuchi K., Mori H., Mori T., Motomura K., Nakade S., Nakamura Y.,
 CC Nishino H., Nishio Y., Oshima T., Saito N., Saito K., Seki Y.,
 CC Sivasubramam S., Tagami H., Takeda J., Takemoto K., Wada C.,
 CC Yamamoto Y., Horikuchi T.;
 CC "A 460-kb DNA sequence of the Escherichia coli K-12 genome
 CC corresponding to the 40.1-50.0 min region on the linkage map".
 RT DNA Res. 3:379-392(1996).
 CC [3]
 CC SEQUENCE FROM N.A.
 CC SPECIES=E.coli; STRAIN=O6:H1 / CFT073 / ATCC 700928;
 RC MEDLINE=22388234; PubMed=12471157;
 RX

RA	Sorochin M., Tascconi E., Takagi T., Takahashi H., Takemaru K.,	CC	
RA	Takuchi M., Tamkooshi A., Tanaka T., Terstutz P., Togomi A.,	CC	
RA	Tosato V., Uchiyama S., Vanderbol M., Vannier F., Vassarotti A.,	CC	
RA	Viari A., Mamout R., Wedler E., Wedler H., Weitzengger T.,	CC	
RA	Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,	CC	
RA	Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.,	CC	
RT	"The complete genome sequence of the Gram-positive bacterium Bacillus	CC	
RT	subtilis".	CC	
RL	Nature 390:249-256 (1997).	CC	
CC	-1- FUNCTION: INVOLVED IN BRANCHED-CHAIN AMINO ACID TRANSPORT.	CC	
CC	-1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).	CC	
CC	-1- SIMILARITY: BELONGS TO THE AZID/H11737/HP1330 FAMILY.	CC	
CC		CC	
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration	CC	
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -	CC	
CC	the European Bioinformatics Institute. There are no restrictions on its	CC	
CC	use by non-profit institutions as long as its content is in no way	CC	
CC	modified and this statement is not removed. Usage by and for commercial	CC	
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/	CC	
CC	or send an email to license@isb-sib.ch).	CC	
CC		CC	
DR	EMBL; Y11043; CAA71994.1; -	CC	
DR	EMBL; U93875; AAB80902.1; -	CC	
DR	EMBL; Z93117; CAB14611.1; -	CC	
DR	PIR; H69592; H69592.	CC	
DR	Subtilist; BG11916; azid.	CC	
DR	Interpro; IPR008407; AZID.	CC	
DR	Pfam; PF05437; Azid; 1.	CC	
KT	Transport; Amino-acid transport; Transmembrane; Complete proteome.	CC	
FT	TRANSMEM 10 30 POTENTIAL.	CC	
FT	TRANSMEM 46 66 POTENTIAL.	CC	
SO	SEQUENCE 110 AA; 11970 MW; 39622669D029033 CRC64;	CC	
		CC	
	Query Match 18.3%; Score 96.5; DB 1; Length 110;	CC	
	Best Local Similarity 27.2%; Pred. No. 0.0059;	CC	
	Matches 31; Conservative 28; Mismatches 42; Indels 13; Gaps 6	CC	
Oy	1 MTTDFSCILIVAVCVITFALRANVPFLIL--KPL-RESQFYCKRMNMNPAGILAILRLA 56		
Db	1 MVTMTQOMITIMVVLGTMLTRFLPFIIPSGKPTPKYVYLDKV--LPSAVIGLVI 57		
Oy	57 STFRSNAIDKTLTFG--LIAAITVAHLIGSRLLSGAGTIVEGVNL 107		
Db	58 YCLKD-VSLISGSHGIPELVGAHAAVLLH-LMKKNMLLSIAGSTVYMYLVOL 108		

RA MEDLINE=92334977; PubMed=1630901;
RA Yura T., Mori H., Nagai H., Nagata T., Ishihama A., Fujita N.,
RA Isono K., Mizobuchi K., Nakara A.,
RA "Systematic sequencing of the Escherichia coli genome: analysis of
RT the 0-2.4 min region.";
RL Nucleic Acids Res. 20:3305-3308(1992).
RN
RN [4]
RN SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden W.A., Rose D.U.,
RA Mau B., Siao Y.,
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
RN
RN [5]
RN SEQUENCE OF 1-28 FROM N.A.
RX MEDLINE=92042080; PubMed=1657980;
RA Karpel R., Alon T., Glaser G., Schuldiner S., Padan E.,
RT "Expression of a sodium proton antiporter (Nhaa) in Escherichia coli
RT is induced by Na⁺ and Li⁺ ions.";
RL J. Biol. Chem. 266:21753-21759(1991).
RN
RN [6]
RN FUNCTION, AND PARTIAL SEQUENCE.
RX MEDLINE=91250446; PubMed=1645730;
RA Taglicht D., Padan E., Schuldiner S.,
RT "Overproduction and purification of a functional Na⁺/H⁺ antiporter
RT coded by nhaa (ant) from Escherichia coli.";
RL J. Biol. Chem. 266:11289-11294(1991).
RN
RN [7]
RN MUTAGENESIS OF HISTIDINE RESIDUES.
RX MEDLINE=93165668; PubMed=8381959;
RA Gerhman Y., Olan Y., Rimon A., Taglicht D., Schuldiner S., Padan E.,
RT "Histidine-226 is part of the pH sensor of Nhaa, a Na⁺/H⁺ antiporter
RT in Escherichia coli.";
RN Proc. Natl. Acad. Sci. U.S.A. 90:1212-1216(1993).
CC
CC -1- FUNCTION: THIS IS A NA(+)/H(+)-ANTIPORTER. IT CAN MEDIATE SODIUM
CC UPTAKE WHEN A TRANSMEMBRANE PH GRADIENT IS APPLIED. IT PROBABLY
CC ACTS IN THE REGULATION OF INTERNAL PH AT THE ALKALINE PH RANGE. IT
CC CATALYZES THE EXCHANGE OF 2H(+) PER NA(+). ITS ACTIVITY IS HIGHLY
CC DEPENDENT ON THE PH.
CC
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane.
CC
CC -1- INDUCTION: TRANSCRIPTION STIMULATED BY HIGH NA⁺ CONCENTRATIONS.
CC
CC This swiss-prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-Brocit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sb.ch/announce/>
CC or send an email to license@isb-sb.ch).
CC
CC -----
DR EMBL: J03879; AAA23448.2; -
DR EMBL: D10483; BAB96582.1; -
DR EMBL: AE000112; AAC73130.1; -
DR EMBL: S67239; AAB20348.1; -
DR PIR: C64722; C64722.
DR BioGene: EG10652; nhaa.
DR InterPro: IP0004670; nhaa.
DR TIGRFAWS; TIGR00773; Nhaa, 1.
KW Transmembrane; Inner membrane; Transport; Antiport; Sodium transport;
KW Complete proteome.
FT TRANSMEM 59 79 POTENTIAL.
FT TRANSMEM 95 115 POTENTIAL.
FT TRANSMEM 125 145 POTENTIAL.
FT TRANSMEM 154 174 POTENTIAL.
FT TRANSMEM 179 199 POTENTIAL.
FT TRANSMEM 205 221 POTENTIAL.
FT TRANSMEM 254 274 POTENTIAL.
FT TRANSMEM 283 299 POTENTIAL.
FT TRANSMEM 328 348 POTENTIAL.
FT TRANSMEM 360 380 POTENTIAL.

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 16, 2004, 06:51:16 ; Search time 6.01671 Seconds
(without alignments)
934.659 Million cell updates/sec

Title: US-10-608-504-5

Perfect score: 526
Sequence: 1 MTDFSCILVAVCAVITF.....RTLLVGAGTIVFGLVNLFP 108

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	96.5	18.3	110	1	AZLD_BACSU
2	81.5	15.5	388	1	NHAA_BACOL
3	80	15.2	628	1	YELF_SCHPO
4	79	15.0	148	1	YEAL_ECOLI
5	75.5	14.4	109	1	YH37_HAEIN
6	75	14.3	1134	1	CVAI_BOVIN
7	74	14.1	171	1	TRPP_CIOAB
8	73.5	14.0	1049	1	ACRB_ECOLI
9	72	13.7	227	1	Y420_NEIMA
10	72	13.7	537	1	OAY_NEUCR
11	70	13.3	342	1	CCR6_CERAE
12	70	13.3	342	1	CCR6_HUMAN
13	70	13.3	342	1	CCR6_PANTR
14	70	13.3	477	1	GTR8_MOUSE
15	70	13.3	652	1	NUSM_PODAN
16	69.5	13.2	521	1	NU2C_SYNY3
17	69.5	13.2	1033	1	ATRA_CANFA
18	69	13.1	342	1	CCR6_MACNE
19	69	13.1	343	1	CCR6_NACPA
20	69	13.1	343	1	CCR6_MACMU
21	69	13.1	462	1	YBXG_BACSU
22	68.5	13.0	507	1	EPOR_MOUSE
23	68	12.9	266	1	UPK_FUSNN
24	68	12.8	714	1	HPPI_AGRTS
25	67.5	12.8	306	1	Y121_WYCGE
26	67.5	12.8	306	1	Y371_MERTJA
27	67.5	12.8	306	1	Y371_MERTJA
28	67.5	12.8	430	1	SECY_STRAU
29	67	12.7	423	1	YHUV_ECOLI
30	67	12.7	715	1	NUSM_NEUCR
31	66.5	12.6	1032	1	ATRA_MOUSE
32	66.5	12.6	1032	1	ATRA_RAT
33	66.5	12.6	1034	1	ATRA_RABIT

34	66	12.5	510	1	D7A1_MOUSE
35	65.5	12.5	199	1	YPHA_BACSU
36	65.5	12.5	250	1	Q8GMS2_synchococ
37	65.5	12.5	277	1	FMF4_ECOLI
38	65.5	12.5	1033	1	ATRA_PIG
39	65.5	12.5	1034	1	ATRA_HUMAN
40	65	12.4	333	1	Y4M0_RHISN
41	65	12.4	414	1	YGJU_ECOLI
42	64.5	12.3	147	1	RLIS_SYNY3
43	64.5	12.3	430	1	SECY_STACA
44	64.5	12.3	457	1	YEBQ_ECOLI
45	64.5	12.3	534	1	HUP3_CHIKE

ALIGNMENTS

RESULT 1
AZLD_BACSU STANDARD; PRT; 110 AA.
ID AZLD_BACSU
AC O07923;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Branched-chain amino acid transport protein azld.
GN AZLD OR BS026700.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
ON NCBI_TaxID=1423;
RX [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168 / BSGC1A1;
RX MEDLINE=97431495; PubMed=9287000;
RA Belitky B.R., Gustafson M.C.U., Sonenshein A.L., von Wachenfeldt C.;
RT "An lrp-like gene of Bacillus subtilis involved in branched-chain
RT amino acid transport.";
RL J. Bacteriol. 179:5448-5457(1997).
[2]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=97453479; PubMed=9308178;
RA Sorokin A., Bolotin A., Purnelle B., Hilbert H., Lauber J.,
RA Duesterhoeft A., Ehrlich S.D.;
RT "Sequence of the Bacillus subtilis genome region in the vicinity of
RT the lrp operon reveals two new extracytoplasmic function RNA
RT polymerase sigma factors Sigv and Sigz.";
RL Microbiology 143:2939-2943(1997).
[3]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=98044033; PubMed=9384377;
RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA Azevedo V., Bertero M.C., Bessieres P., Bolotin A., Borzhet S.,
RA Bottaris R., Boursier L., Brans A., Braun M., Brimell S.C., Bron S.,
RA Brouillet S., Brusch C.V., Caldwell B., Capuano V., Carter N.M.,
RA Choi S.K., Codani J.J., Conerton I.F., Cummings N.D., Daniel R.A.,
RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
RA Enlian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
RA Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grand G.,
RA Guisepi G., Guy B.J., Haga K., Hatach J., Harwood C.R., Hanaut A.,
RA Hilbert H., Holstappel S., Hosono S., Hullo M.F., Iteya M., Jones L.,
RA Joris B., Karamata D., Kasahara Y., Kleier-Blanchard W., Klein C.,
RA Kobayashi Y., Koester P., Koningsstein G., Krogh S., Kumano M.,
RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
RA Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue V.,
RA Medina N., Mellado R.P., Mizuno M., Mostl D., Nakai S., Noback M.,
RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudaga B., Park S.H.,
RA Parro V., Pohl T.M., Portelle B., Rapoport G., Rey W., Reynolds S.,
RA Preechan E., Pujic P., Purnelle B., Rapoport G., Rey W., Reynolds S.,
RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadle Y.,
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
RA Sekiguchi J., Sekowska A., Seror S.J., Seror P., Shin B.S., Soldo B.,

submitted to the EMBL Data Library, July 1999

A:Description: *Pyrococcus abyssi* genome sequence: insights into archaeal chromosome stru
A:Reference number: A75001
A:Accession: C75003
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-314 <KAM>
A:Cross-references: GB:A248288; GB:AL096836; NID:95458960; PIDN:CAB50465.1; PID:el51636
A:Experimental source: Strain Orsay
C:Genetics:
A:Gene: PAH1029

Query Match 15.0%; Score 79; DB 2; Length 314;
Best Local Similarity 26.7%; Pred. No. 1.4;
Matches 20; Conservative 17; Mismatches 26; Indels 12; Gaps 1;

QY 4 DFGCILVAVCAVITPALRAVPFLIKPLRESQFVGKAMMPAGIATITASTFRSNA 63
DB 238 DIIDLMVIGASIT-----RPIKVDPSIVKVTPIVLVMAITVSLFRNNK 285

QY 64 IDLKITLFTGLAVAI 78
DB 286 VGRKTAATLLLVSTI 300

RESULT 14

AB1068
Probable membrane protein STY4872 [imported] - *Salmonella enterica* subsp. *enterica* serov
C:Species: *Salmonella enterica* subsp. *enterica* serovar Typh
A>Note: This species has also been called *Salmonella typh*
C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
C:Accession: AB1068
R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Main, J.; Churcher,
th, T.; Connor, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
, S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
A:Title: Complete genome sequence of a multiple drug resistant *Salmonella enterica* serov
A:Reference number: AB0502; MUID:21534947; PMID:11677608
A:Accession: AB1068
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-394 <PAR>
A:Cross-references: GB:AL513382; PIDN:CAD03361.1; PID:G16505633; GSPDB:GN00176
C:Genetics:
A:Gene: STY4872

Query Match 15.0%; Score 79; DB 2; Length 394;
Best Local Similarity 26.2%; Pred. No. 1.7;
Matches 27; Conservative 25; Mismatches 43; Indels 8; Gaps 4;

QY 9 LLVAVCAVITPALRAVPFLIKPLR-ESQFVGKAMMPAGIATITASTFRSNAIDLK 67
DB 261 ILIGACVLLSLASDSILLIISIGFATFMTGTSIVMP--LARQLSAPGNITLGLIV 337

QY 68 TITFGLIAVITVAHLG-GRRITLSV---GAGTIVFVGLVN 106
DB 338 TITVGIQITLGPLASLGSNGASAIINATLGCAPALFPALIS 380

RESULT 15

A99631
Na+/H antiporter [imported] - *Escherichia coli* (strain O157:H7, substrain RIMD 050952)

C:Species: *Escherichia coli*
C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 03-Aug-2001
C:Accession: A99631

R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
gaawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shingawa, H.
DNA Res. 8, 11-22, 2001

A:Title: Complete genome sequence of enterohemorrhagic *Escherichia coli* O157:H7 and geno
A:Reference number: A99629; MUID:21156231; PMID:11258796

A:Accession: A99631
A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-388 <HAY>
A:Cross-references: GB:BA000007; PIDN:BA833440.1; PID:G13359473; GSPDB:GN00154
A:Experimental source: Strain O157:H7, substrain RIMD 050952
C:Genetics:
A:Gene: Ecs0017
C:Superfamily: Na+/H+-exchanging protein nhaA

Query Match 14.9%; Score 78.5; DB 2; Length 388;
Best Local Similarity 30.3%; Pred. No. 1.9;
Matches 27; Conservative 18; Mismatches 35; Indels 9; Gaps 3;

QY 18 ITPALRAVPFLIK-PLRESQFVGKAMMPAGIATITASTFRSNAIDKITLFTGLIAY 76
DB 134 IAFNLGVLLLSRVPLAKIFL--MALAIDDLGAIITIALFYTNDSLMSLGVAAVAI 191

QY 77 AITVAHLGGRITLSVAGTIVFVGLV 105
DB 192 AVLAVALNLGCVART-----GVYILVGVV 214

Search completed: April 16, 2004, 06:54:26
Job time : 10.4234 secs

Query Match	15.0%	Score 79;	DB 2;	Length 148;
Best Local Similarity	22.5%;	Pred. No. 0.71;		
Matches	27;	Conservative	20;	Mismatches 49; Indels 24; Gaps 2
QY	12	VAVCAVTFPLRANPFLINLPKPRSCQVGMAMMPAGIIATLTAFTFSNAIDLTKLTFE	71	
DB	24	VASIVLPIIVRYPTLPSTFFPMIEKQSLSTCIILITGVAPLPSGLPSTLIHSPILNW	83	
QY	72	-GLIATAITVAHLIGSRFTLL-----SAGAGIVVGVVNL	107	
DB	84	KSLVAIVAGIVSWLGRGVTLMGSPOLVAGLIVGVTLGVALFRGVVPGVGLIAGVSL	143	

Query Match	15.0%;	Score 79;	DB 2;	Length 148;
Best Local Similarity	22.5%;	Pred. No. 0.71;		
Matches	27;	Conservative	20;	Mismatches 49; Indels 24; Gaps 2
Qy	12	VAVCAITATTAATAPVLLILKLPEESQFVGMAMMDAGILAITASTFRSNAIDLKLTLE	71	
Db	24	VASIIILIVIKTPELSTPFFMEIKQSLGIIILITGVAPLASGLPSTLIHSPLNW	83	
Qy	72	GLIAVAITYVAALLGGRTLL-----SQAAGTIYFVGIVNL	107	
Db	84	KSIIAIVAAVGIVSWLGGRTVTLGSGQPOLVAGILVGTGVALFRGVVPVPLITGIVSL	144	

```
RESULT 13
C75003
na/cas: exchanging protein related PAB1029 - Pyrococcus abyssi (strain Orsay)
C:Species: Pyrococcus abyssi
C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Aug-1999
C:Accession: C75003
R:anonymous, Genoscope
```

```

Query Match      15.5% Score 81.5; DB 2; Length 388;
Best Local Similarity 30.3%; Pred. No. 0.96;
Matches 27; Conservative 18; Mismatches 35; Indels 9; Gaps 3;

QY    18 TFPALRAVDFLLIK-PLRESQFVGKMMMPGAILALTASTFRSNALDITKLTKTGLIAY 76
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB     134 IFALGVALLTLSSRVPALKRIPL--MALAIIDDLGAIIIIILFYNDLSMASLGVAAVAI 191
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY     77 AITVVAAHLGGRRRTLLSVAGTIIVEGYLV 105
      | : | : | | | | | | | | | | | | | | | | | | | | | | | | | |
DB     192 AVLAVALNLGARRT-----GVYLIVGV 214

RESULT 7
C95963
conserved hypothetical membrane protein SMD21531 (imported) - Sinorhizobium meliloti (S)
C:Species: Sinorhizobium meliloti
C:Date: 24-Aug-2001 #sequence revision 24-Aug-2001 #text change 30-Sep-2001

```

```

Qy      8  ILVAVCAVTFELKLAPEFLIKLPLEESQFVGWMMMPAGLIAITTAFTFRS----- 61
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      231 VTMLAPVLIPLISLVRSP-----QEGASIGKRAPLVEGFVLAFLVLAFGNSAGLVPP 283

Qy      62 -----NAIDKITFTFGIAIVATVVAHILGRRITLTSVCAGTIVFGLVNLFP 108
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      284 IASEVMAISRMWALLAGVAVGKMT-----SLRRVLEVGGDVAVALVATLTF 330

RESULT 8
PB2726
AAC transporter ATP-binding protein XF1061 [imported] - Xylella fastidiosa (strain 9a5c)
C:Species: Xylella fastidiosa
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
C:Accession: PB2726
R:Anonymus, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequencing
A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A:Reference number: A82515; MUID:20365717; PMID:10910347
A:Note: for a complete list of authors see reference number A59328 below
A:Accession: PB2726
A:Status: preliminary
A:Molecule type: DNA
A:Accession: 1-589 <SW>
A:Cross-references: GB:AEO003944; GB:AEO003849; NID:G9106023; PIDN:AA893891.1; GSPDB:GN000
A:Experimental source: strain 9a5c
R:Simson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.;
B:Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carrero, D.M.; Carret,
As-Nero, E.; Docena, C.; El-Dorry, H.; Facinanti, A.P.; Ferreira, A.J.S.
A:Submitted to GenBank, June 2000
A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohn
J.D.; Jungueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Klegger, J.E.; Kuramae, E.E.; Laig
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins,
A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;
F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.;
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.B.; de Sa, R.G.; Santelli, R.V.; Sawada
A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silve
M.; Tesharo, M.H.; Vallada, H.; Van Sluys, M.A.; Verjowski-Almeida, S.; Vettore, A.L.;
A:Reference number: A59328
A:Contents: annotation
C:Genetics:
A:Gene: XF1061

Query Match      15.2%;      Score 80.      DB 2;      Length 589;
Best Local Similarity      34.7%;      Pred. No. 2;

```

Db 58 YCLMD-VSLSGSGHICFELVGAARVVLH-LMKKNMLLSLAGGVVYVMVVL 108

RESULT 2

D86826
hypothetical protein yqfC [imported] - Lactococcus lactis subsp. lactis (strain IL1403)
C/Species: Lactococcus lactis subsp. lactis
C/Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 03-Aug-2001
C/Accession: D86826
R/Author: A.; Winkler, P.; Manger, S.; Jallion, O.; Malarme, K.; Weisenbach, J.; Ehrlich
Genome Res. 11, 731-753, 2001
A/Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis
A/Reference number: A86625; MUID:21235186; PMID:11337471
A/Accession: D86826
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-108 <STO>
A/Cross-references: GB:AE005176; PID:g12724619; PID:AAK05710.1; GSPDB:GN00146
A/Experimental source: strain IL1403
C/Genetics:
A/Gene: yqfC

Query Match 18.0%; Score 94.5; DB 2; Length 108;
Best Local Similarity 29.9%; Pred. No. 0.015;
Matches 32; Conservative 24; Mismatches 42; Indels 9; Gaps 4;

Qy 3 TDFSCILVAVCAVITFALRAVPELILKPLRESQFVGKMMMPAGILAITAST-FRS 61
Db 2 SSEEFSLITIGCAITWISRVLPFLTKKMSLPQVVEYLSFVPAVIMISALMISNLFQ 61

Qy 62 NAIDKLTTF-GLIAVITVAHLGGRRTLSVGAGTIVFGVNL 107
Db 62 HLGHLPSVWNNMLASIPVLAITL--TKNLL-----VVLGVPSL 101

RESULT 3

AH1617
hypothetical protein lml1481 [imported] - Listeria innocua (strain C11p11262)
C/Species: Listeria innocua
C/Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001
C/Accession: AH1617
R/Glaser, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker
D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.
Science 294, 849-852, 2001
A/Authors: Kreft, U.; Kuhn, M.; Kunst, F.; Kurapkac, G.; Madueno, E.; Matcournam, A.; Me
ok, C.; Schluter, T.; Simoes, N.; Tixeret, A.; Vazquez-Boland, J.A.; Voss, H.; Weiland,
A/Title: Comparative genomics of Listeria species.
A/Reference number: AB1077; MUID:21537279; PMID:11679669
A/Accession: AH1617
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-108 <GLA>
A/Cross-references: GB:AL592022; PIDN:CA096712.1; PID:g16413954; GSPDB:GN00178
A/Experimental source: strain C11p11262
C/Genetics:
A/Gene: lml1481

Query Match 16.1%; Score 84.5; DB 2; Length 108;
Best Local Similarity 24.8%; Pred. No. 0.15; Indels 9; Gaps 3;
Matches 25; Conservative 24; Mismatches 43;

Qy 9 LTVAVCAVITFALRAVPELILKPLRESQFVGKMMMPAGILAITASTFRSNAIDK 67
Db 8 LVLVIGGLVTFPRVLPFIFVRKLPQSVVIRVLSVPLCITLALFVQSLITNENSP 67

Qy 68 TLTF-GLIAVITVAHLGGRRTLSVGAGTIVFGVNL 107
Db 68 SINENLALPLTIT-----AILTKNLMWIVIGIISM 101

RESULT 4

Db7534
hypothetical protein CC2301 [imported] - Caulobacter crescentus
C/Species: Caulobacter crescentus
C/Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 20-Apr-2001
C/Accession: D87534
R/Nerman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
B.; Laub, M.T.; Deboy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kojic
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A/Title: Complete Genome Sequence of Caulobacter crescentus.
A/Reference number: A87249; MUID:21173698; PMID:11259647
A/Accession: D87534
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-407 <STO>
A/Cross-references: GB:AE005673; NID:g13423820; PIDN:AAK24272.1; GSPDB:GN00148
C/Genetics:
A/Gene: CC2301

Query Match 15.9%; Score 83.5; DB 2; Length 407;
Best Local Similarity 30.0%; Pred. No. 0.63; Indels 19; Gaps 5;
Matches 33; Conservative 20; Mismatches 38;

Qy 8 ILTVAVCAVITFA---LRAVPELIL-----KPLRESQFVGKMM-MMPAGIL-----AI 53
Db 130 MLVGAVALGLITRADGVLKAIYOVNLNLEERIRDLAMVKVSMATYPAILLFWCVRI 189

Qy 54 LTASTFRSNAIDKLTTFGLIAVITVAHLGGRRTLSVGAGTIVFG 103
Db 190 LDANNFRG-----RNIAIALIAVCIYAGATLGADAPIALILIGVAMIG 234

RESULT 5

AC1255
hypothetical protein lmo1443 [imported] - Listeria monocytogenes (strain EGD-e)
C/Species: Listeria monocytogenes
C/Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001
C/Accession: AC1255
R/Glaser, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker
D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.
Science 294, 849-852, 2001
A/Authors: Kreft, U.; Kuhn, M.; Kunst, F.; Kurapkac, G.; Madueno, E.; Matcournam, A.; Me
ok, C.; Schluter, T.; Simoes, N.; Tixeret, A.; Vazquez-Boland, J.A.; Voss, H.; Weiland,
A/Title: Comparative genomics of Listeria species.
A/Reference number: AB1077; MUID:21537279; PMID:11679669
A/Accession: AC1255
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-108 <GLA>
A/Cross-references: GB:NC 003210; PIDN:CA099521.1; PID:g16410872; GSPDB:GN00177
A/Experimental source: strain EGD-e
C/Genetics:
A/Gene: lmo1443

Query Match 15.8%; Score 83; DB 2; Length 108;
Best Local Similarity 26.9%; Pred. No. 0.21; Indels 10; Gaps 3;
Matches 29; Conservative 23; Mismatches 46;

Qy 6 SCILVAVCAVITFALRAVPELILKPLRESQFVGKMMMPAGILAITASTFRSNAID 65
Db 5 SYTLFVIGGLVTFPRVLPFIFVRKLPQSVVIRVLSVPLCITLALFVQSLITNENSP 61

Qy 66 LKTLTFG-----LIAVITVAHLGGRRTLSVGAGTIVFGVNL 108
Db 62 RENSPPGINENLNASLPTITAIL--TKNLMWIVVIGIISMALIRLP 107

RESULT 6

CE4722
Na+/H+-exchanging protein nhaA - Escherichia coli (strain K-12)
C/Species: Escherichia coli

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OK protein - protein search, using sw model

Run on: April 16, 2004, 06:51:17 ; Search time 8.4234 Seconds
(without alignments)
1233.313 Million cell updates/sec

Title: US-10-608-504-5

Perfect score: 526

Sequence: 1 MTTDFSCILVAVAVCAVITF.....RTLLSVGNGTIVFGLVNL 108

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

PIR_78: *
1: p1r1: *
2: p1r2: *
3: p1r3: *
4: p1r4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	96.5	18.3	110	1 H69592	branched-chain ami
2	94.5	18.0	108	2 D86826	hypothetical prote
3	84.5	16.1	108	2 AH1617	hypothetical prote
4	83.5	15.9	407	2 D87534	hypothetical prote
5	83	15.8	108	2 AC1255	hypothetical prote
6	81.5	15.5	388	2 C64722	Na+/H+-exchanging
7	80.5	15.3	345	2 C95963	conserved hypothet
8	80	15.2	589	2 F82726	ABC transporter At
9	80	15.2	628	2 T38034	probable transport
10	79	15.0	148	2 E64939	hypothetical prote
11	79	15.0	148	2 B90941	hypothetical prote
12	79	15.0	148	2 F85789	hypothetical prote
13	79	15.0	314	2 C75003	na+/ca+ exchanging
14	79	15.0	394	2 AB1068	probable membrane
15	78.5	14.9	388	2 A99631	Na+/H antiporter
16	78.5	14.9	388	2 A85482	cytochrome-c oxida
17	77.5	14.7	235	2 D48366	hypothetical prote
18	76	14.4	102	2 T35134	glutathione-regula
19	76	14.4	656	2 B82056	branched-chain ami
20	75.5	14.4	182	1 B64041	hypothetical prote
21	75.5	14.4	182	2 F72623	conserved hypothet
22	75	14.3	923	2 AC2705	hypothetical prote
23	75	14.3	923	2 C97487	adenylate cyclase
24	74.5	14.2	379	2 A41350	Gung protein - Xan
25	74.5	14.2	379	2 S67856	uncharacterized me
26	74	14.1	171	2 A97344	sugar ABC transport
27	74	14.1	324	2 A83940	probable Na+/Ca2+-
28	74	14.1	324	2 D71159	ATPase related to
29	73.5	14.0	443	2 H96939	

30	73.5	14.0	1049	2 B36938	acetylflavin resist
31	73.5	14.0	1049	2 C90693	acridine efflux pu
32	73.5	14.0	1049	2 G85543	acridine efflux pu
33	73	13.9	405	2 A12934	chromate transport
34	73	13.9	409	2 F98347	probable transport
35	72.5	13.8	156	2 B84227	hypothetical prote
36	72.5	13.8	461	2 E81862	probable integral
37	72.5	13.8	461	2 B81080	drug resistance tr
38	72	13.7	227	2 G81014	conserved hypothet
39	72	13.7	537	2 G31277	guinate transport
40	71.5	13.6	101	2 AF2830	conserved hypothet
41	71.5	13.6	104	2 C97608	hypothetical prote
42	71.5	13.6	286	2 C71059	hypothetical prote
43	71.5	13.6	297	2 A83049	hypothetical prote
44	71.5	13.6	351	2 T44798	iron transport mem
45	71.5	13.6	395	2 A82827	conserved hypothet

ALIGNMENTS

RESULT 1

branched-chain amino acid transport protein azid - Bacillus subtilis

C/Species: Bacillus subtilis

C/Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 28-Jul-2003

C/Accession: H69592; T44778

R/Kunst, F.; Ogatawara, N.; Moser, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Beres, C.; Bron, S.; Brouillet, S.; Brunsch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Chk

A.; Ehrlich, S.D.; Emmerson, P.T.; Entlin, K.D.; Errington, J.; Fabret, C.; Ferrari, E.

Nature 390:249-256, 1997

A/Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gallier

lech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hultio, M.F.

Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,

A/Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Maguda, S.; Mauesel

Y, M.; Ogawa, K.; Ogiwara, A.; Oudeg, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle

Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadate, Y.; Sato, T.; Scanlon,

A/Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seriz

akeuchi, M.; Yamakoshi, A.; Tanaka, T.; Terpetra, P.; Tognoni, A.; Tosato, V.; Uchiyama,

T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, I

A/Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Darchin, A.

A/Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.

A/Reference number: A69580; MUID:98044033; PMID:9384377

A/Accession: H69592

A/Status: nucleic acid sequence not shown; translation not shown

A/Molecule type: DNA

A/Residues: 1-110 <KUN>

A/Cross-references: GB:299117; GB:AL009126; NID:G2634966; PIDN:CAB14611.1; PID:G2635115

A/Experimental source: strain 168

J.Belitsky, B.R.; Gustafsson, M.C.U.; Sonenshein, A.L.; von Wachenfeldt, C.

U. Bacteriol. 179: 5448-5457, 1997

A/Title: An lrp-like gene of Bacillus subtilis involved in branched-chain amino acid tra

A/Reference number: Z2837; MUID:97431495; PMID:9287000

A/Accession: T44778

A/Status: translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-110 <BEL>

A/Cross-references: EMBL:Y11043; NID:G1926275; PIDN:CAA71941.1; PID:G1926282

A/Experimental source: strain 1A1

C/Genetics:

A/Genes: azid

C/Superfamily: branched-chain amino acid transport protein, Azid type

C/Keywords: branched-chain amino acid transport; transmembrane protein

Query Match 18.3%; Score 96.5; DB 1; Length 110;

Best local similarity 27.2%; Pred. No. 0.0095; Indels 13; Gaps 6;

Matches 31; Conservative 26; Mismatches 42;

Qy 1 MTTDFSCILVAVAVCAVITFALRAVPLIL--KPL-RESQFYGRKMMMPAGILAILTA 56

Db 1 MTTMTQOMITIAVAVGTMLTRFLPMPISGKPTPKYQYGVKVV--LPSAVIGLVI 57

Qy 57 STFSNMIIDKTLTFG---LIIVAITVVAHLGSRITLLSVGNGTIVFGLVNL 107

QY 92 LSVGAGTIVF-----VGLVNL 108
DB 163 IAGIGIIFGFGNGDAIGLSNLM 187

RESULT 12

US-10-767-701-40503
Sequence 40503, Application US/10767701
GENERAL INFORMATION:
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(5353)B
CURRENT FILING DATE: 2004-01-29
NUMBER OF SEQ ID NOS: 63128
SBO ID NO 40503
LENGTH: 119
TYPE: PRT
ORGANISM: Sorghum bicolor
FEATURE:
OTHER INFORMATION: Clone ID: SORBI-28MAY03-C65105_1.pep
US-10-767-701-40503

Query Match 13.0%; Score 68.5; DB 6; Length 119;
Best Local Similarity 30.8%; Pred. No. 0.57; Indels 33; Gaps 6;
Matches 33; Conservative 15; Mismatches 26; Indels 33; Gaps 6;

QY 8 ILVVAVCAVITFA--LRAVPLILKPLRESQVGRKAMMPAGILALITASTFRSNAID 65
DB 18 ILSVIVAVITFSGMGAIPTLMNSEI-----LPSI-----KSLGGS 56

QY 66 LKTLFGLIAVATVVAHL-----GGRITLS---VGAGTIVFVGL 104
DB 57 IATLANWLTSAITMTNIMLTWSVG--TFLSYWVSAFTLVFVL 101

RESULT 13

US-60-556-841-2298
Sequence 2298, Application US/60556841
GENERAL INFORMATION:
APPLICANT: Adad, Mark S.
TITLE OF INVENTION: Genes and Uses for Plant Improvement
FILE REFERENCE: 38-21(53450)
CURRENT APPLICATION NUMBER: US/60/556,841
CURRENT FILING DATE: 2004-03-25
NUMBER OF SEQ ID NOS: 12463
SEQ ID NO 2298
LENGTH: 475
TYPE: PRT
ORGANISM: Pseudomonas putida KT2440
US-60-556-841-2298

Query Match 13.0%; Score 68.5; DB 7; Length 475;
Best Local Similarity 25.9%; Pred. No. 2.9; Indels 25; Gaps 7;
Matches 30; Conservative 24; Mismatches 37; Indels 25; Gaps 7;

QY 15 CAVITPALRAVPLILKPLRESQVGRKAMMPAGILALITASTFRSNAI-----DL 66
DB 131 CA--SLIWSVHVLVAGIKAFINQVTVAKVPLLFVIAAFAPFADLFTDINWL 188

QY 67 KTLTFLI-----AVATV-----AHLIGRITLS-VGAGTIV-FVGLVNL 107
DB 189 SNPOFSVLEQVNMMLVTVFVFIGIGASVYSGRAQRSDVGAIVIGFLVAL 244

RESULT 14
US-10-779-543-23512
Sequence 23512, Application US/10779543
GENERAL INFORMATION:
APPLICANT: Williams et al

TITLE OF INVENTION: GENE PRODUCTS DIFFERENTIALLY EXPRESSED
FILE REFERENCE: IN CANCEROUS CELLS AND THEIR METHODS OF USE II
CURRENT APPLICATION NUMBER: US/10/779,543
CURRENT FILING DATE: 2004-02-12
PRIOR APPLICATION NUMBER: 10/076,555
PRIOR FILING DATE: 2002-02-15
PRIOR APPLICATION NUMBER: 09/217,471
PRIOR FILING DATE: 1998-12-21
PRIOR APPLICATION NUMBER: 60/068,755
PRIOR FILING DATE: 1997-12-23
PRIOR APPLICATION NUMBER: 60/080,664
PRIOR FILING DATE: 1998-04-03
PRIOR APPLICATION NUMBER: 60/105,234
PRIOR FILING DATE: 1998-10-21
PRIOR APPLICATION NUMBER: 09/297,648
PRIOR FILING DATE: 2000-04-10
PRIOR APPLICATION NUMBER: PCT/US99/01619
PRIOR FILING DATE: 1999-01-28
PRIOR APPLICATION NUMBER: 60/072,910
PRIOR FILING DATE: 1998-01-28
PRIOR APPLICATION NUMBER: 60/075,954
PRIOR FILING DATE: 1998-02-24
PRIOR APPLICATION NUMBER: 60/080,114
PRIOR FILING DATE: 1998-03-31
Remaining Prior Application data removed - See File Wrapper or PALM.

NUMBER OF SEQ ID NOS: 23767
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 23512
LENGTH: 537
TYPE: PRT
ORGANISM: Homo sapiens
US-10-779-543-23512

Query Match 12.8%; Score 68; DB 6; Length 537;
Best Local Similarity 34.9%; Pred. No. 3.8; Indels 8; Gaps 3;
Matches 22; Conservative 13; Mismatches 20; Indels 8; Gaps 3;

QY 53 ILTASTFRSN-AIDLKTLFGLIAVATVVAHL-----GGRITLSVG---ACTIVFVGL 104
DB 224 VLKSVGRQSNMAASLSTGVVAVKSTIPATLIVHVSKEPLICGSSVMAASLVTMGI 283

QY 105 VNL 107
DB 284 VNL 286

RESULT 15

US-10-767-701-43433
Sequence 43433, Application US/10767701
GENERAL INFORMATION:
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(5353)B
CURRENT APPLICATION NUMBER: US/10/767,701
CURRENT FILING DATE: 2004-01-29
NUMBER OF SEQ ID NOS: 63128
SEQ ID NO 43433
LENGTH: 119
TYPE: PRT
ORGANISM: Sorghum bicolor
FEATURE:
OTHER INFORMATION: Clone ID: SORBI-28MAY03-C71332_1.pep
US-10-767-701-43433

Query Match 12.8%; Score 67.5; DB 6; Length 119;
Best Local Similarity 26.5%; Pred. No. 0.74; Indels 27; Gaps 3;
Matches 26; Conservative 15; Mismatches 30; Indels 27; Gaps 3;

QY 11 VVAVCAVITFAVPLILKPLRESQVGRKAMMPAGILALITASTFRSNAIDLKTLT 70


```

APPLICANT: Wendy S. Halsey
APPLICANT: Jon Chambers
APPLICANT: Allison Isobel Muir
APPLICANT: Philip Graham Szekeres
APPLICANT: Usman Shadon
APPLICANT: Derek J. Bergsma
APPLICANT: Nabil A. Eshourbagy
APPLICANT: David Michalovich
APPLICANT: Pamela A. Lane
APPLICANT: Menelas N. Pangalos
APPLICANT: Melanie Robbins
APPLICANT: David Malcolm Duckworth
APPLICANT: Jeffrey Hill
APPLICANT: Ping Tsui
APPLICANT: Rankeji Agarwal
APPLICANT: Randall Forrest Smith
APPLICANT: Lisa Vawter
APPLICANT: Catherine E. Ellis
APPLICANT: Manhandeswar Gattu
APPLICANT: John W. Quillen, Jr.
APPLICANT: Erin M. Roland
APPLICANT: Steven Michael Foord
APPLICANT: Han Ngoc Trinh
APPLICANT: Alexander Taylor
APPLICANT: Henry Sarau
APPLICANT: Steven Ruben
APPLICANT: George H. Poste
APPLICANT: Michel Louis Souchet
APPLICANT: Philippe Laurent Robert
APPLICANT: Stephane Clement Krief
APPLICANT: Bernard Emile Joseph Gout
APPLICANT: Eve Mahe
TITLE OF INVENTION: HUMAN G PROTEIN-COUPLED RECEPTOR
TITLE OF INVENTION: POLYPEPTIDES AND POLYNUCLEOTIDES
PTE REFERENCE: GP-70775B-C1
CURRENT FILING DATE: 2003-09-22
PRIOR APPLICATION NUMBER: US/10/667,762
PRIOR FILING DATE: 2003-04-07,386
PRIOR APPLICATION NUMBER: US/10/407,386
PRIOR FILING DATE: 2003-04-04
PRIOR APPLICATION NUMBER: US/10/278,107
PRIOR FILING DATE: 2002-10-21
PRIOR APPLICATION NUMBER: US//10/125,749
PRIOR FILING DATE: 2002-04-19
PRIOR APPLICATION NUMBER: US/09/988,922
PRIOR FILING DATE: 2001-11-19
NUMBER OF SEQ ID NOS: 82
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 76
LENGTH: 342
TYPE: PRT
ORGANISM: HOMO SAPIENS
US-10-667-762-76

Query Match 13.3%; Score 70; DB 6; Length 342;
Best Local Similarity 27.8%; Pred. No. 1.3;
Matches 25; Conservative 12; Mismatches 37; Indels 16; Gaps 2

QY 7 CILVVAVCVITFPALRAV-----PFLIKPFEESQFVGKMMNPAGILAI 53
   :|::|::|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
DB 36 CMLIVVFVCGVENSLSLVYVLSIFPKKLSINDVPLVNLPLADLVFVCTLPFMAVAGIHEM 95
   :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|

QY 54 LTASTFRSNAIDIKTLTF--GLIIVATV 80
   :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
DB 96 VFGQVWCKSLGIYITNFYTSMLIDTCITV 125
   :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|

RESULT 10
Sequence 9, Application US/10787018
GENERAL INFORMATION:
APPLICANT: Goelling, Jennifa
APPLICANT: Daireghil, Daniel J.
APPLICANT: Hanley, Michael

```

```

APPLICANT: Miao, Zhenhua
APPLICANT: Talbot, Dale
APPLICANT: Schall, Thomas J.
APPLICANT: Chemocentryx, Inc.
TITLE OF INVENTION: Chemokine Receptor
FILE REFERENCE: 019934-0007210US
CURRENT FILING DATE: 2004-02-24
PRIORITY APPLICATION NUMBER: US/09/721,495
PRIORITY FILING DATE: 2000-11-21
PRIORITY APPLICATION NUMBER: US 60/159, 015
PRIORITY FILING DATE: 1999-10-12
PRIORITY APPLICATION NUMBER: US 60/159, 210
PRIORITY FILING DATE: 1999-10-13
PRIORITY APPLICATION NUMBER: US 60/172, 979
PRIORITY FILING DATE: 1999-12-20
PRIORITY APPLICATION NUMBER: US 60/173, 388
PRIORITY FILING DATE: 1999-12-28
PRIORITY APPLICATION NUMBER: US 60/186, 626
PRIORITY FILING DATE: 2000-03-03
PRIORITY APPLICATION NUMBER: US 09/666, 019
PRIORITY FILING DATE: 2000-10-10
NUMBER OF SEQ ID NOS: 14
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 9
LENGTH: 342
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: chemokine receptor STRL33
US-10-787-018-9

Query Match          13.3%; Score 70; DB 6; Length 342;
Best Local Similarity 27.8%; Pred. No. 1.3;
Matches 25; Conservative 12; Mismatches 37; Indels 16; Gaps 2;

QY      7 CILVVAVCAVITFALAAV-----PRLIKPLRESQFVKMMAMPAGILAI 53
DB      36 CMYVVVVCGLVGNLSLVISIFYHKQSLTDVFLVNLPLADLVCTCTPMAVAGIHEW 95
          :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
          54 LTASTFRSNAIDKLTLT---GLIAVAIV 80
DB      96 VFGQWCKSLIGYITNYTSMILTCITV 125
          :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :

RESULT 11
US-60-556-841-1117
US-Sequence 1117, Application US/60556841
GENERAL INFORMATION:
APPLICANT: Abad, Mark S.
TITLE OF INVENTION: Genes and Uses for Plant Improvement
FILE REFERENCE: 38-21(53450)
CURRENT APPLICATION NUMBER: US/60/556,841
CURRENT FILING DATE: 2004-03-25
NUMBER OF SEQ ID NOS: 12463
SEQ ID NO 1117
LENGTH: 462
TYPE: PRT
ORGANISM: Bacillus subtilis subsp. subtilis str. 168
US-60-556-841-1117

Query Match          13.1%; Score 69; DB 7; Length 462;
Best Local Similarity 22.1%; Pred. No. 2.4;
Matches 32; Conservative 24; Mismatches 45; Indels 44; Gaps 7

QY      8 ILVVAVCAV-ITFALRAV-PRLIKPLRES-----QFVKMA-----MMPDAG 49
DB      43 VLVAVAGIPIFIMRAMEGMLYVESGTSFATFGHQYIHMPAGYITAMSNMFGWITV 102
          :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
          50 ILAIIKASTFRSNAI-DKLTIFGLIAVAIVVAVHLIGGR-----TL 91
DB      103 MSEIIANGSYTKYMPDPLPANI PGIVAVVILGAANLISVKSFGFEFFFMATIKVTIILM 162
          :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :

```

TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 HYPOTHEICAL: YES
 ORIGINAL SOURCE:
 ORGANISM: Enterococcus faecium
 FEATURE:
 NAME/KEY: misc feature
 LOCATION: (B) LOCATION 1...326
 SEQUENCE DESCRIPTION: SEQ ID NO: 3665:
 US-10-417-884A-3665

Query Match 13.8%; Score 72.5; DB 6; Length 326;
 Best Local Similarity 27.4%; Pred. No. 0.64;
 Matches 29; Conservative 16; Mismatches 38; Indels 23; Gaps 4;

QY 8 ILVV---AVCAVTFPALRAVPFLILKPLEESQFVG-KAMMMFAGILAITASTRSN 62
 DB 50 ILVTRLPRTICLLIVGATSSICGLIMQHLTONKFSPTAGTWSATGILVAMIFLPG 109
 QY 63 AIDKTLTFGLIAVAITVAHLGGRRTLLSVGAGTIVFVGLVNL 108
 DB 110 A-----SLVRSLTAFCPAF-----AGTLFLSLTRLF 137

RESULT 7

US-10-417-884A-5593
 Sequence 5593, Application US/10417884A

GENERAL INFORMATION:
 APPLICANT: Lynn A Doucette-Stamm and David Bush
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
 ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
 NUMBER OF SEQUENCES: 7310
 CORRESPONDENCE ADDRESS:
 ADDRESS: GENOME THERAPEUTICS CORPORATION
 STREET: 100 Beaver Street
 CITY: Waltham
 STATE: Massachusetts
 COUNTRY: USA
 ZIP: 02354

COMPUTER READABLE FORM:
 MEDIUM TYPE: CD-ROM ISO9660

OPERATING SYSTEM: <Unknown>
 SOFTWARE: ASCII

CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/10/417,884A
 FILING DATE: 17-Apr-2003
 PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/09/107,532
 FILING DATE: 30-Jun-1998
 APPLICATION NUMBER: 60/085,598
 FILING DATE: 14 May 1998
 APPLICATION NUMBER: 60/051571
 FILING DATE: July 2, 1997

ATTORNEY/AGENT INFORMATION:
 NAME: Arinello, Pamela Deneke
 REGISTRATION NUMBER: 40,489
 REFERENCE/DOCKET NUMBER: GTC-012
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (781)893-5007
 TELEFAX: (781)893-8277

INFORMATION FOR SEQ ID NO: 5593:
 SEQUENCE CHARACTERISTICS:

LENGTH: 117 amino acids
 TYPE: amino acid

TOPOLOGY: linear
 MOLECULE TYPE: protein
 HYPOTHEICAL: YES
 ORIGINAL SOURCE:

ORGANISM: Enterococcus faecium
 FEATURE:
 NAME/KEY: misc_feature

LOCATION: (B) LOCATION 1...117
 SEQUENCE DESCRIPTION: SEQ ID NO: 5593:
 US-10-417-884A-5593

Query Match 13.7%; Score 72; DB 6; Length 117;
 Best Local Similarity 25.7%; Pred. No. 0.22;
 Matches 27; Conservative 24; Mismatches 50; Indels 4; Gaps 3;

QY 5 FSCILVAVCAVTFPALRAVPFLILKPLEESQFVGKAMMMFAGILAITASTRSNA 63
 DB 14 FETTLTITIGCTLATLSVLPFLKFDLPQPLLEYLSFVPIYMSALMFSSTFONI 73
 QY 64 IDKTLTF-GLIAVAITVAHLGGRRTLLSVGAGTIVFVGLVNL 107
 DB 74 GHLPOINTENALASVPTLAAILS--KSLVIVLAGILSLIRL 116

RESULT 8

US-09-804-291A-449
 Sequence 449, Application US/09804291A

GENERAL INFORMATION:
 APPLICANT: ZOZULA, SERGEY

TITLE OF INVENTION: HUMAN OLFACTORY RECEPTORS AND GENES ENCODING SAME
 FILE REFERENCE: 100337,54287US

CURRENT APPLICATION NUMBER: US/09/804,291A
 CURRENT FILING DATE: 2001-03-13

PRIOR APPLICATION NUMBER: 60/188,914
 PRIOR FILING DATE: 2000-03-13

PRIOR APPLICATION NUMBER: 60/192,033
 PRIOR FILING DATE: 2000-03-24

PRIOR APPLICATION NUMBER: 60/198,474
 PRIOR FILING DATE: 2000-04-14

PRIOR APPLICATION NUMBER: 60/199,335
 PRIOR FILING DATE: 2000-04-24

PRIOR APPLICATION NUMBER: 60/207,702
 PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: 60/213,849
 PRIOR FILING DATE: 2000-06-23

PRIOR APPLICATION NUMBER: 60/226,534
 PRIOR FILING DATE: 2000-08-16

PRIOR APPLICATION NUMBER: 60/230,732
 PRIOR FILING DATE: 2000-09-07

PRIOR APPLICATION NUMBER: 60/266,862
 PRIOR FILING DATE: 2001-02-07

NUMBER OF SEQ ID NOS: 531
 SOFTWARE: PatentIn Ver. 3.2

SEQ ID NO: 449
 LENGTH: 313

TYPE: PRT
 ORGANISM: Homo sapiens

US-09-804-291A-449

Query Match 13.5%; Score 71; DB 5; Length 313;
 Best Local Similarity 26.4%; Pred. No. 0.91;
 Matches 29; Conservative 18; Mismatches 41; Indels 22; Gaps 3;

QY 8 ILVAVCAVTFPALRAVPFLILKPLEESQFVGK---MAMMMFAGILAITASTRSNAI 64
 DB 106 IVLGISCCILTMAYRYAICPLKYSTLSPRACTANVGSSWLTGITATTHS--- 162
 QY 65 DKTLTFGLIAVAITVAHL-----GGRRTLLSVGAGTIVFV 102
 DB 163 ----LIFSLPFRSHPIIFHLCDILPVLRLASAKRHSISVWTATVFI 208

RESULT 9

US-10-667-762-76

Sequence 76, Application US/10667762
 GENERAL INFORMATION:

APPLICANT: Erding Hu
 APPLICANT: Yuan Zhu
 APPLICANT: Ganesh M. Sathe

APPLICANT: Joyce Yue Mao

Tue Apr 20 06:47:21 2004

us-10-608-504-5.rapn

Page 2

SEQ ID NO 8672
LENGTH: 568
TYPE: PRT
ORGANISM: Zea mays
US-60-556-841-8672

Query Match 13.9%; Score 73; DB 7; Length 568;
Best Local Similarity 31.3%; Pred. No. 1.1;
Matches 36; Conservative 17; Mismatches 44; Indels 18; Gaps 7;

QY 6 SCILLVAVCAVITFALRAVPELIL---KPLRESQFVGK-----MAMMPAGILAILTAST 58
DB 290 ACVLLVAVCAVLTIT---GMVPYTLGSDAPLAEA-FAAKGLKFTVTLISGAVAGLTITL 345

QY 59 FRSNAIDKLTITFGL-IAVAITVAHLGGRTLLS---VGAQTIFFVGLVNL 107
DB 346 LVGLVYQSR-LYFGLGRDGLPSVFAVHPTHTPVQSQIIVGCVAAVAGLFTNV 399

RESULT 3
US-60-556-841-11808
Sequence 11808, Application US/60556841
GENERAL INFORMATION:
APPLICANT: Abad, Mark S.
TITLE OF INVENTION: Genes and Uses for Plant Improvement
FILE REFERENCE: 38-21(53450)
CURRENT APPLICATION NUMBER: US/60/556,841
CURRENT FILING DATE: 2004-03-25
NUMBER OF SEQ ID NOS: 12463
SEQ ID NO 11808
LENGTH: 568
TYPE: PRT
ORGANISM: Zea mays
US-60-556-841-11808

Query Match 13.9%; Score 73; DB 7; Length 568;
Best Local Similarity 31.3%; Pred. No. 1.1;
Matches 36; Conservative 17; Mismatches 44; Indels 18; Gaps 7;

QY 6 SCILLVAVCAVITFALRAVPELIL---KPLRESQFVGK-----MAMMPAGILAILTAST 58
DB 290 ACVLLVAVCAVLTIT---GMVPYTLGSDAPLAEA-FAAKGLKFTVTLISGAVAGLTITL 345

QY 59 FRSNAIDKLTITFGL-IAVAITVAHLGGRTLLS---VGAQTIFFVGLVNL 107
DB 346 LVGLVYQSR-LYFGLGRDGLPSVFAVHPTHTPVQSQIIVGCVAAVAGLFTNV 399

RESULT 4
US-60-556-841-9476
Sequence 9476, Application US/60556841
GENERAL INFORMATION:
APPLICANT: Abad, Mark S.
TITLE OF INVENTION: Genes and Uses for Plant Improvement
FILE REFERENCE: 38-21(53450)
CURRENT APPLICATION NUMBER: US/60/556,841
CURRENT FILING DATE: 2004-03-25
NUMBER OF SEQ ID NOS: 12463
SEQ ID NO 9476
LENGTH: 618
TYPE: PRT
ORGANISM: Oryza sativa
US-60-556-841-9476

Query Match 13.9%; Score 73; DB 7; Length 618;
Best Local Similarity 29.1%; Pred. No. 1.2;
Matches 34; Conservative 17; Mismatches 44; Indels 22; Gaps 6;

QY 6 SCILLVAVCAVITFALRAVPELIL---KPLRESQFVGK-----MAMMPAGILAIL 54
DB 321 ACVLLVAVCAVLTIT---GMVPYTLGSDAPLAEA-FAAKGLKFTVTLISGAVAGLTITL 376

QY 55 TASTFRSNAIDKLTITFGLIAVAITVAHLGGRTLLS---VGAQTIFFVGLVNL 107

DB 377 LVGLVYQSR-LYFGLGRDGLP---SIFAKVHPTHTPILHSQIIVGCVAAVAGLFTNV 430

RESULT 5
US-60-556-841-11757
Sequence 11757, Application US/60556841
GENERAL INFORMATION:
APPLICANT: Abad, Mark S.
TITLE OF INVENTION: Genes and Uses for Plant Improvement
FILE REFERENCE: 38-21(53450)
CURRENT APPLICATION NUMBER: US/60/556,841
CURRENT FILING DATE: 2004-03-25
NUMBER OF SEQ ID NOS: 12463
SEQ ID NO 11757
LENGTH: 618
TYPE: PRT
ORGANISM: Oryza sativa
US-60-556-841-11757

Query Match 13.9%; Score 73; DB 7; Length 618;
Best Local Similarity 29.1%; Pred. No. 1.2;
Matches 34; Conservative 17; Mismatches 44; Indels 22; Gaps 6;

QY 6 SCILLVAVCAVITFALRAVPELIL---KPLRESQFVGK-----MAMMPAGILAIL 54
DB 321 ACVLLVAVCAVLTIT---GMVPYTLGSDAPLAEA-FAAKGLKFTVTLISGAVAGLTITL 376

QY 55 TASTFRSNAIDKLTITFGLIAVAITVAHLGGRTLLS---VGAQTIFFVGLVNL 107
DB 377 LVGLVYQSR-LYFGLGRDGLP---SIFAKVHPTHTPILHSQIIVGCVAAVAGLFTNV 430

RESULT 6
US-10-417-884A-3665
Sequence 3665, Application US/10417884A
GENERAL INFORMATION:
APPLICANT: Lynn A Doucette-Stamm and David Bush
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
NUMBER OF SEQUENCES: 7310
CORRESPONDENCE ADDRESS:
ADDRESSEE: GENOME THERAPEUTICS CORPORATION
STREET: 100 Beaver Street
CITY: Waltham
STATE: Massachusetts
COUNTRY: USA
ZIP: 02354
COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660
COMPUTER: PC
OPERATING SYSTEM: <Unknown>
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/417,884A
FILING DATE: 17-Apr-2003
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/107,532
FILING DATE: 30-Jun-1998
APPLICATION NUMBER: 60/085,598
FILING DATE: 14 May 1998
APPLICATION NUMBER: 60/051571
FILING DATE: July 2, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Ariniello, Pamela Deneke
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-012
TELEPHONE: (781) 893-5007
TELEFAX: (781) 893-8277
INFORMATION FOR SEQ ID NO: 3665
SEQUENCE CHARACTERISTICS:
LENGTH: 326 amino acids

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 16, 2004, 06:53:52 ; Search time 7.52089 Seconds
(without alignments)
386.360 Million cell updates/sec

Title: US-10-608-504-5

Perfect score: 526
Sequence: 1 MTDFSCILVAVCAVITF.....RTLSVGAGTIVGVGLNLF 108

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 193692 seqs, 26905285 residues

Total number of hits satisfying chosen parameters: 193692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents AA New:*
1: /cgn2_6/ptodata/1/paa/PCIT NEW COMB.pep:*
2: /cgn2_6/ptodata/1/paa/US06 NEW COMB.pep:*
3: /cgn2_6/ptodata/1/paa/US07 NEW COMB.pep:*
4: /cgn2_6/ptodata/1/paa/US08 NEW COMB.pep:*
5: /cgn2_6/ptodata/1/paa/US09 NEW COMB.pep:*
6: /cgn2_6/ptodata/1/paa/US10 NEW COMB.pep:*
7: /cgn2_6/ptodata/1/paa/US60 NEW COMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	82.5	15.7	156	US-10-724-972A-6988	Sequence 6988, Ap
2	73	13.9	568	US-60-556-841-8672	Sequence 8672, Ap
3	73	13.9	568	US-60-556-841-11808	Sequence 11808, A
4	73	13.9	618	US-60-556-841-9476	Sequence 9476, Ap
5	73	13.9	618	US-60-556-841-11757	Sequence 11757, A
6	72.5	13.8	326	US-10-417-884A-3665	Sequence 3665, Ap
7	72	13.7	117	US-10-417-884A-5593	Sequence 5593, Ap
8	71	13.5	313	US-09-804-291A-449	Sequence 449, App
9	70	13.3	342	US-10-667-762-76	Sequence 76, Appl
10	70	13.3	342	US-10-787-018-9	Sequence 9, Appl
11	69	13.1	462	US-60-556-841-11117	Sequence 11117, A
12	68.5	13.0	119	US-10-767-701-40503	Sequence 40503, A
13	68.5	13.0	475	US-60-556-841-2298	Sequence 2298, Ap
14	68	12.9	537	US-10-779-543-23512	Sequence 23512, A
15	67.5	12.8	119	US-10-767-701-43433	Sequence 43433, A
16	67.5	12.8	432	US-60-556-841-6678	Sequence 6678, Ap
17	67.5	12.8	432	US-60-556-841-6915	Sequence 6915, Ap
18	67	12.7	174	US-60-556-841-8566	Sequence 8566, Ap
19	67	12.7	300	US-60-556-841-11087	Sequence 11087, A
20	67	12.7	447	US-60-556-841-3680	Sequence 3680, Ap
21	67	12.7	447	US-60-556-841-10069	Sequence 10069, A
22	66.5	12.6	398	US-10-724-972A-4715	Sequence 4715, Ap
23	65.5	12.5	351	US-10-767-701-46482	Sequence 46482, A
24	65.5	12.5	351	US-10-724-972A-4781	Sequence 4781, Ap
25	65.5	12.5	440	US-10-724-972A-6421	Sequence 6421, Ap
26	65.5	12.5	486	US-60-556-841-4550	Sequence 4550, Ap

27	65	12.4	129	US-10-767-701-54323	Sequence 54323, A
28	65	12.4	465	US-10-417-884A-5620	Sequence 5620, Ap
29	65	12.4	469	US-60-556-841-4458	Sequence 4458, Ap
30	65	12.4	507	US-60-556-841-4463	Sequence 4463, A
31	65	12.4	508	US-60-556-841-12239	Sequence 12239, A
32	65	12.4	509	US-60-556-841-3141	Sequence 3141, A
33	64.5	12.3	181	US-60-556-841-940	Sequence 940, App
34	64.5	12.3	181	US-60-556-841-9000	Sequence 9000, Ap
35	64.5	12.3	423	US-10-724-972A-4088	Sequence 4088, Ap
36	64.5	12.3	605	US-10-724-972A-4854	Sequence 4854, Ap
37	64	12.2	458	US-10-767-701-52684	Sequence 52684, A
38	64	12.2	425	US-60-556-841-56380	Sequence 56380, A
39	63.5	12.1	297	US-10-724-972A-5519	Sequence 5519, A
40	63.5	12.1	311	US-10-767-701-43663	Sequence 43663, A
41	63.5	12.1	348	US-10-417-884A-6950	Sequence 6950, Ap
42	63.5	12.1	502	US-10-724-972A-7473	Sequence 7473, Ap
43	63.5	12.1	507	PCT-US04-09416-2	Sequence 2, Appl
44	63.5	12.1	518	US-10-724-972A-6261	Sequence 6261, Ap
45	63.5	12.1	572	US-60-556-841-7889	Sequence 7889, Ap

ALIGNMENTS

```
RESULT 1
US-10-724-972A-6988
Sequence 6988, Application US/10724972A
GENERAL INFORMATION:
APPLICANT: Douncette-Stamm, Lynn
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
FILE REFERENCE: PATR03-16
CURRENT APPLICATION NUMBER: US/10/724,972A
PRIOR FILING DATE: 2003-12-01
PRIOR APPLICATION NUMBER: 09/450,969
PRIOR FILING DATE: 1999-11-29
PRIOR APPLICATION NUMBER: 09/134,001
PRIOR FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: 60/064,964
PRIOR FILING DATE: 1997-11-08
PRIOR APPLICATION NUMBER: 60/055,779
PRIOR FILING DATE: 1997-08-14
NUMBER OF SEQ ID NOS: 7544
SEQ ID NO 6988
LENGTH: 156
TYPE: PRT
ORGANISM: S.epidermidis
US-10-724-972A-6988
Query Match      15.7%; Score 82.5; DB 6; Length 156;
Best Local Similarity 25.7%; Pred. No. 0.019;
Matches 28; Conservative 22; Mismatches 44; Indels 15; Gaps 5;
Cy      1 MTDFSCILVAVCAVITFALRAVPFLIKPRSEQFVGKMMMPAGIATITAFSTR 60
      |||:::||||:|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      48 MTDLHLVIV-LGVVTLVIVIPIMSRVNLPAIVKMSFIP---ITLFTLII- 102
      |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
      61 SNAIDLTFTFG-----LIAVAIVVAHLGGRRTLSVGAGTIVFV 102
      |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      103 DGVIGQDHAFGYTLNPLPIAIVFVMAIFRSLVITLIGG---IFV 148
      |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
RESULT 2
US-60-556-841-8672
Sequence 8672, Application US/60556841
GENERAL INFORMATION:
APPLICANT: Abad, Mark S.
TITLE OF INVENTION: Genes and Uses for Plant Improvement
FILE REFERENCE: 38-21(53450)
CURRENT APPLICATION NUMBER: US/60/556,841
PRIOR FILING DATE: 2004-03-25
NUMBER OF SEQ ID NOS: 12463
```

PRIOR APPLICATION NUMBER: US 60/128,706
 PRIOR FILING DATE: 1999-04-09
 NUMBER OF SEQ ID NOS: 8344
 SEQ ID NO 5990
 LENGTH: 494
 TYPE: PRT
 ORGANISM: *Proteus mirabilis*
 US-09-543-681A-5990

Query Match 13.6%; Score 71.5; DB 4; Length 494;
 Best Local Similarity 25.4%; Pred. No. 1.2;
 Matches 33; Conservative 25; Mismatches 43; Indels 29; Gaps 6;

QY 8 ILVVAVCAVITFALRAVPLILKPLRESQFVGK-MAMMPAG-----LALITASGF- 59
 DB 250 IIAVVLGGLPYGIRIFDIKLVWCKMMLAKIHLIGLPRAGENLWILHTASAFI 309
 QY 60 ---RSNAIDKLTLPGL-----IAVAI---TVVAHLGGR-----RTLASVAGT 98
 DB 310 GLMGVPLAAGTLYFQLSLFMLPGIATISGNEILVGHVGAKKFDPAFTGFKSLKAGV 369
 QY 99 IYFVGLVNL 108
 DB 370 LFTIGVVIY 379

RESULT 13
 US-09-328-352-6931
 Sequence 6931, Application US/09328352
 Patent No. 6562958
 GENERAL INFORMATION:
 APPLICANT: Gary L. Breton et al.
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
 FILE REFERENCE: GTC99-03PA
 CURRENT APPLICATION NUMBER: US/09/328,352
 CURRENT FILING DATE: 1999-06-04
 NUMBER OF SEQ ID NOS: 8252
 SEQ ID NO 6931
 LENGTH: 169
 TYPE: PRT
 ORGANISM: *Acinetobacter baumannii*
 US-09-328-352-6931

Query Match 13.5%; Score 71; DB 4; Length 169;
 Best Local Similarity 20.4%; Pred. No. 0.35;
 Matches 28; Conservative 25; Mismatches 48; Indels 36; Gaps 4;

QY 8 ILVVAVCAVITFALRAVPLILKPLRESQFVGK-MAMMPAGILALIT 55
 DB 30 VLLVLLICGLSONAAVITAGILIVIKITPLNPFPIYQAGLNLGIIILITGVTPIA 89
 QY 56 ASTFRSNAIDKLTLP-GLIYAVITVAHLGGR-----RTL 91
 DB 90 SGLTSGESILKSPISFKSLVATAIGLVAMLGSRGVKMSQPDVVAHLIGTVAGVALL 149
 QY 92 LSYGAGTIVFVGLVNL 108
 DB 150 RGVVPGLIAGLSLF 166

RESULT 14
 US-09-489-039A-7704
 Sequence 7704, Application US/09489039A
 Patent No. 6610836
 GENERAL INFORMATION:
 APPLICANT: Gary Breton et al.
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
 FILE REFERENCE: 2709.2004001
 CURRENT APPLICATION NUMBER: US/09/489,039A
 CURRENT FILING DATE: 2000-01-27
 PRIOR APPLICATION NUMBER: US 60/117,747

PRIOR FILING DATE: 1999-01-29
 NUMBER OF SEQ ID NOS: 14342
 SEQ ID NO 7704
 LENGTH: 188
 TYPE: PRT
 ORGANISM: *Klebsiella pneumoniae*
 US-09-489-039A-7704

Query Match 13.5%; Score 71; DB 4; Length 188;
 Best Local Similarity 19.8%; Pred. No. 0.4;
 Matches 24; Conservative 23; Mismatches 50; Indels 24; Gaps 2;

QY 12 VAVCAVITFALRAVPLILKPLRESQFVGK-MAMMPAGILALITASTFRSNAIDKLTLP 71
 DB 64 VASILVLIIVRTPLNAFFPWKEGTLVGIILITGVMAPIASCTLPSTLIHSFNMV 123
 QY 72 -GLIYAVITVAHLGGRRTL-----SYGAGTIVFVGLVNL 107
 DB 124 KSLALAVGVFVWSLGRGVSLMSQPHLVAGLVGTVLGVALLFRGVVPGLIAGTISL 183
 QY 108 F 108
 DB 184 F 184

RESULT 15
 US-09-328-352-6237
 Sequence 6237, Application US/09328352
 Patent No. 6562958
 GENERAL INFORMATION:
 APPLICANT: Gary L. Breton et al.
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
 FILE REFERENCE: GTC99-03PA
 CURRENT APPLICATION NUMBER: US/09/328,352
 CURRENT FILING DATE: 1999-06-04
 NUMBER OF SEQ ID NOS: 8252
 SEQ ID NO 6237
 LENGTH: 259
 TYPE: PRT
 ORGANISM: *Acinetobacter baumannii*
 US-09-328-352-6237

Query Match 13.4%; Score 70.5; DB 4; Length 259;
 Best Local Similarity 22.4%; Pred. No. 0.69;
 Matches 38; Conservative 24; Mismatches 37; Indels 71; Gaps 8;

QY 8 ILVVAVCAVITFALRAVPLILKPLRESQFVG 40
 DB 78 LNAVIAICALISSPAGACSLIPQSVLRPFVVMILVIAITLVKKNFGQVHTGKIT 137
 QY 41 KAMMPAGILALITASTFRSNAIDKLTLP-----DXTLT 70
 DB 138 KMLVLAGISLAFYDGFPGTGFPIRFLQVDFLHASLSKIGNFTNLALS 197
 QY 71 F-----GLIYAVITVAHLGGRRTLTVGAG--TIVFVGLVNL 107
 DB 198 FLIPTGAILIHLGMLMAANVLGSIV-GVNTALKYSGSFRIIFLIVSI 246

Search completed: April 16, 2004, 06:56:35
 Job time : 11.2284 secs

/ APPLICANT: Gary Breton et. al
 / TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
 / FILE REFERENCE: 2709.2004001
 / CURRENT APPLICATION NUMBER: US/09/489,039A
 / PRIOR FILING DATE: 2000-01-27
 / PRIOR APPLICATION NUMBER: US 60/117,747
 / PRIOR FILING DATE: 1999-01-29
 / NUMBER OF SEQ ID NOS: 14342
 / SEQ ID NO 13014
 / LENGTH: 1049
 / TYPE: PRT
 / ORGANISM: Klebsiella pneumoniae
 / US-09-489-039A-13014

Query Match 13.8%; Score 72.5; DB 4; Length 1049;
 Best Local Similarity 19.4%; Pred. No. 2.2; Mismatches 32; Indels 53; Gaps 3;

QY 5 FSCILIVAVCAVITFALRAVPLIKPLRESQFVGKAMMPAGILAIITSTR--- 60
 DB 877 YALSLIVFCLALAYESWSIPFSVW-----LVVPLGVGALAAATRGULTN 923
 QY 61 -----SNAI-----DKTLTFGLIAVITVAHL 84
 DB 924 DYFQVGLTTTGLSKAMILIVEPAKDMEXGKGLIATLEAVRMRLPILMTSLAFI 983
 QY 85 LGGRRTLSVAGT 98
 DB 984 LGVMPVIVISSAGS 997

RESULT 10
 US-09-107-532A-5593
 / Sequence 5593, Application US/09107532A
 / Patent No. 6583275
 / GENERAL INFORMATION:
 / APPLICANT: Lynn A Doucette-Stamm and David Bush
 / TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
 / ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
 / NUMBER OF SEQUENCES: 7310
 / CORRESPONDENCE ADDRESS:
 / ADDRESS: GENOME THERAPEUTICS CORPORATION
 / STREET: 100 Beaver Street
 / CITY: Waltham
 / STATE: Massachusetts
 / COUNTRY: USA
 / ZIP: 02354
 / COMPUTER READABLE FORM:
 / MEDIUM TYPE: CD-ROM ISO9660
 / COMPUTER: PC
 / OPERATING SYSTEM: <Unknown>
 / SOFTWARE: ASCII
 / CURRENT APPLICATION DATA:
 / APPLICATION NUMBER: US/09/107,532A
 / FILING DATE: 30-Jun-1998
 / PRIOR APPLICATION DATA:
 / APPLICATION NUMBER: 60/085,598
 / FILING DATE: 14 May 1998
 / APPLICATION NUMBER: 60/051,571
 / FILING DATE: July 2, 1997
 / ATTORNEY/AGENT INFORMATION:
 / NAME: Arianello, Pamela Deneka
 / REGISTRATION NUMBER: 40,489
 / REFERENCE/DOCKET NUMBER: GTC-012
 / TELECOMMUNICATION INFORMATION:
 / TELEPHONE: (781) 893-5007
 / TELEFAX: (781) 893-8277
 / INFORMATION FOR SEQ ID NO: 5593:
 / SEQUENCE CHARACTERISTICS:
 / LENGTH: 117 amino acids
 / TYPE: amino acid
 / TOPOLOGY: linear

/ MOLECULE TYPE: protein
 / HYPOTHETICAL: YES
 / ORIGINAL SOURCE:
 / ORGANISM: Enterococcus faecium
 / FEATURE:
 / NAME/KEY: misc feature
 / LOCATION: (B) LOCATION 1...117
 / SEQUENCE DESCRIPTION: SEQ ID NO: 5593:
 / US-09-107-532A-5593

Query Match 13.7%; Score 72; DB 4; Length 117;
 Best Local Similarity 25.7%; Pred. No. 0.16; Indels 4; Gaps 3;

QY 5 FSCILIVAVCAVITFALRAVPLIKPLRESQFVGKAMMPAGILAIITSTRSNA 63
 DB 14 FEFLLITLIGCTLATMLSRVLPVLLKFPDLQPLEVLSFVIVIMSALMTSLSFTONI 73
 QY 64 IDKTLTF-GLIAVITVAHLGRRITLSVAGTIVGVNL 107
 DB 74 GHLPLINENALASVPTLLAITS--KSLIVLVLAGILSLIRL 116

RESULT 11
 US-09-252-991A-27266
 / Sequence 27266, Application US/09252991A
 / Patent No. 6551795
 / GENERAL INFORMATION:
 / APPLICANT: Marc J. Rubenfield et al.
 / TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 / TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
 / FILE REFERENCE: 107196,136
 / CURRENT APPLICATION NUMBER: US/09/252,991A
 / CURRENT FILING DATE: 1999-02-18
 / PRIOR APPLICATION NUMBER: US 60/074,788
 / PRIOR FILING DATE: 1998-02-18
 / PRIOR APPLICATION NUMBER: US 60/094,190
 / PRIOR FILING DATE: 1998-07-27
 / NUMBER OF SEQ ID NOS: 33142
 / SEQ ID NO 27266
 / LENGTH: 303
 / TYPE: PRT
 / ORGANISM: Pseudomonas aeruginosa
 / US-09-252-991A-27266

Query Match 13.6%; Score 71.5; DB 4; Length 303;
 Best Local Similarity 27.4%; Pred. No. 0.63; Indels 43; Gaps 7;

QY 4 DFSCILIVAVCAVITFALRAVPLIK-----PLRESQFVG-----KXAMN 45
 DB 35 DFPLPLF-----CA-LRFALALPLVLRGMPAPFGRIILIGVLGVKFGLLFVGAAQG 89
 QY 46 MEAGILAIITASTRNSAIDKTLTFGLIAVITVAHLGRRITLSVAGTIV 100
 DB 90 MPAGLSLVLSQ-----VFETIIIAIMLGERPSPGRLGSLAAGLV 134
 QY 101 FVGL 104
 DB 135 LIGL 138

RESULT 12
 US-09-543-681A-5990
 / Sequence 5990, Application US/09543681A
 / Patent No. 6605709
 / GENERAL INFORMATION:
 / APPLICANT: GARY BRETON
 / TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
 / TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
 / FILE REFERENCE: 2709.1002-001
 / CURRENT APPLICATION NUMBER: US/09/543,681A
 / CURRENT FILING DATE: 2000-04-05

Db 186 AIGGLVVAASHLLVATLTPAKRPRLWTTGNAALLFLG-VNYY 229

RESULT 6
US-09-245-039-2

; Sequence 2, Application US/09245039
; Patent No. 6555522
; GENERAL INFORMATION:
; APPLICANT: Iyengar, Srinivas R.V.
; APPLICANT: Meng, Gezhi
; APPLICANT: Chen, Yibang
; APPLICANT: Weinstein, Harel
; APPLICANT: Buck, Elizabeth
; TITLE OF INVENTION: PEPTIDES AND OTHER SMALL MOLECULES DERIVED FROM REGIONS
; TITLE OF INVENTION: OF INTERACTING PROTEINS AND USES THEREOF
; FILE REFERENCE: 6923-074
; CURRENT APPLICATION NUMBER: US/09/245,039
; CURRENT FILING DATE: 1999-02-05
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1134
; TYPE: PRT
; ORGANISM: Bos taurus
US-09-245-039-2

Query Match 14.3%; Score 75; DB 4; Length 1134;

Best Local Similarity 28.6%; Pred. No. 1.2; 45; Indels 10; Gaps 5;
Matches 30; Conservative 20; Mismatches

QY 12 VAVCAVITFALRAVPEFLIKPLRESQFVGKXAMMPAGILALTASTRSNAI----DLK 67

DB 127 LALLSLTFPALLCCFPALGGPAGAGAAVATADQGVWQLLV-TVSYALLPVRSL 185

QY 68 TLTPGLIAVA--ITVAHLGGRRLT--SVGAGTIVVGLVNL 108

DB 186 AIGGLVVAASHLLVATLTPAKRPRLWTTGNAALLFLG-VNYY 229

RESULT 7

US-09-489-039A-7412
; Sequence 7412, Application US/09489039A

; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 7412
; LENGTH: 590
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-7412

Query Match 13.9%; Score 73; DB 4; Length 590;

Best Local Similarity 25.6%; Pred. No. 0.94;
Matches 32; Conservative 19; Mismatches 40; Indels 34; Gaps 7;

QY 10 LVAVCAVITFALRAVP--FLIKPLRESQFVGKXAMMPAGILALTASTR---RSNA 63

DB 16 VVVCITAVLY---VPLFTIVQSPFLSAPFSRSKWSLEAFLETPDPDYLAIRSGF 71

QY 64 IDLKTTLFGILAAVAT--VVAHL-----LGGRTT-----LSVAGTIVFVG 103

DB 72 I-----LAFGLVIAITFLGILALMTWRTDLPGRRIIEPLIVPIFVSPVNLGGVYVAA 127

QY 104 LVNL 108

DB 128 PVGFF 132

RESULT 8
US-09-107-532A-3665

; Sequence 3665, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESS: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCIT
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 3665:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 326 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...326
; SEQUENCE DESCRIPTION: SEQ ID NO: 3665:
US-09-107-532A-3665

Query Match 13.8%; Score 72.5; DB 4; Length 326;

Best Local Similarity 27.4%; Pred. No. 0.52;
Matches 29; Conservative 16; Mismatches 38; Indels 23; Gaps 4;

QY 8 ILIVV---AVCAVITFALRAVPEFLIKPLRESQFVG-KXAMMPAGILALTASTRSN 62

DB 50 ILVTRLPRLICLIVATSSICGLIMQHUTONKFVSPPTAGTWDASRLILVAMTLP 109

QY 63 AIDKTLTFLGILAAVATVVAHLGGRRLTSLVGAGTIVFVGLVNL 108

DB 110 A-----SLIVSLTAFCFAF-----AGTLFLSTLRLE 137

RESULT 9
US-09-489-039A-13014

; Sequence 13014, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:

Tue Apr 20 06:47:21 2004

us-10-608-504-5.rat

Page 2

CURRENT FILING DATE: 2000-04-05
PRIOR APPLICATION NUMBER: US 60/128,706
PRIOR FILING DATE: 1999-04-09
NUMBER OF SEQ ID NOS: 8344
SEQ ID NO 6109
LENGTH: 405
TYPE: PRT
ORGANISM: Proteus mirabilis
US-09-543-681A-6109

Query Match
Best Local Similarity 31.7%; Pred. No. 0.0012;
Matches 33; Conservative 18; Mismatches 30; Indels 23; Gaps 6;

QY 11 VAVCAVITPALRAVPFLIKPLRE-SQVGVKAMMPAGIL-ALITSTRSNAT 64
DB 292 IAGCFIVTLLAAV-----REISWFIGLYFLW---GILGATPVTALISRTAE 340

QY 65 DKTITFGLIAVATVVAHLGRRLLSVGAGTIVGVLNLF 108
DB 341 SKQGYILGLVQ-SISQFASWVG---IALGGILVWPGTILTF 378

RESULT 3
US-09-134-001C-5172
Sequence 5172, Application US/09134001C
Patent No. 6389370
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
FILE REFERENCE: GTC-007
CURRENT APPLICATION NUMBER: US/09/134,001C
CURRENT FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/064,964
PRIOR FILING DATE: 1997-11-08
PRIOR APPLICATION NUMBER: US 60/055,779
PRIOR FILING DATE: 1997-08-14
NUMBER OF SEQ ID NOS: 5674
SEQ ID NO 5172
LENGTH: 156
TYPE: PRT
ORGANISM: Staphylococcus epidermidis
US-09-134-001C-5172

Query Match
Best Local Similarity 25.7%; Pred. No. 0.012;
Matches 28; Conservative 22; Mismatches 44; Indels 15; Gaps 5;

QY 1 MTTFSCITLIVAVCAVITPALRAVPFLIKPLRESQVGVKAMMPAGILALITSTR 60
DB 48 MTTFDLITLIV-LGCVVTLVLRVPIRIMSRVLPALVTKWLSFIP---ITLFTALIT- 102
QY 61 SNMIDLTTFG-----LIAVAITVVAHLGRRLLSVGAGTIVFV 102
DB 103 DGVIGCHDAFGYTLNPIYIAIVPYVMAIFRSLTITLIGS---IFV 148

RESULT 4
US-09-540-236-2360
Sequence 2360, Application US/09540236
Patent No. 6673910
GENERAL INFORMATION:
APPLICANT: Gary L. Bretton et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATAR
FILE REFERENCE: 2709.2005-001
CURRENT APPLICATION NUMBER: US/09/540,236
CURRENT FILING DATE: 2000-04-04
NUMBER OF SEQ ID NOS: 3840
SEQ ID NO 2360
LENGTH: 506
TYPE: PRT

ORGANISM: M. catarrhalis
US-09-540-236-2360

Query Match
Best Local Similarity 14.3%; Score 75; DB 4; Length 506;
Matches 28; Conservative 19; Mismatches 46; Indels 28; Gaps 3;

QY 11 VAVCAVITPALRAVP-----LILKPLRESQVGVKAMMP 46
DB 224 VVMTLAAAFKLSAAPHWSVDYTGAPPAIPLASVSKVAMMALRPFVGSYFTM 283
QY 47 PAG---LIALITSTRSNATDKTITFGLIAVATVVAHLGRRLLSVGAGTIVFG 103
DB 284 PAFETLIVIVLSILGNLAIKRONNIKRM-LAYSSIAHIGVALTALISGAGSLPVS 342
QY 104 L 104
DB 343 M 343

RESULT 5
US-08-726-214-2
Sequence 2, Application US/08726214
Patent No. 6107076
GENERAL INFORMATION:
APPLICANT: Tang, Wei-Jen
APPLICANT: Gilman, Alfred G.
TITLE OF INVENTION: SOLUBLE MAMMALIAN ADENYLYL CYCLASE
TITLE OF INVENTION: AND USES THEREFOR
NUMBER OF SEQUENCES: 31
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: Texas
COUNTRY: United States of America
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/726,214
FILING DATE: Concurrently Herewith
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/005,498
FILING DATE: 04-OCT-1995
ATTORNEY/AGENT INFORMATION:
NAME: Highlander, Steven L.
REGISTRATION NUMBER: 37,642
REFERENCE/DOCKET NUMBER: UTSD:450
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (512) 474-7577
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1134 amino acids
TYPE: amino acid
STRANDEDNESS: linear
TOPOLOGY: linear
US-08-726-214-2

Query Match
Best Local Similarity 14.3%; Score 75; DB 3; Length 1134;
Matches 30; Conservative 20; Mismatches 45; Indels 10; Gaps 5;

QY 12 VAVCAVITPALRAVPFLIKPLRESQVGVKAMMPAGILALITSTRSNAT---DLK 67
DB 127 LALPLSLTFALCCFPALGPGAGAHAGAAVATADQGWQLLIV-IFVSYALLPVRSLL 185
QY 68 TLITFGLIAV--ITVVAHLGRRLL--SVGAGTIVGVLNLF 108

Tue Apr 20 06:47:21 2004

US-10-608-504-5.ra1

Page 1

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 16, 2004, 06:51:17 ; Search time 10.2284 Seconds
(without alignments)
545.110 Million cell updates/sec

Title: US-10-608-504-5

Perfect score: 526
Sequence: 1 MTFPSCILLVAVCAVITF.....RTLSVGAGTIVFGLVNL 108

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database:

Issued Patents RA:
1: /cgn2_6/prodata/2/1aa/5A.COMB.pep:*
2: /cgn2_6/prodata/2/1aa/5B.COMB.pep:*
3: /cgn2_6/prodata/2/1aa/6A.COMB.pep:*
4: /cgn2_6/prodata/2/1aa/6B.COMB.pep:*
5: /cgn2_6/prodata/2/1aa/PCTUS.COMB.pep:*
6: /cgn2_6/prodata/2/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	526	100.0	108	4	US-09-471-803A-5
2	94.5	18.0	405	4	US-09-543-681A-6109
3	82.5	15.7	156	4	US-09-134-001C-5172
4	75	14.3	506	4	US-09-540-236-2360
5	75	14.3	1134	3	US-08-726-214-2
6	75	14.3	1134	4	US-09-245-039-2
7	73	13.9	590	4	US-09-489-039A-7412
8	72.5	13.8	326	4	US-09-107-532A-3665
9	72.5	13.8	1049	4	US-09-489-039A-13014
10	72	13.7	117	4	US-09-107-532A-5593
11	71.5	13.6	303	4	US-09-252-991A-21266
12	71.5	13.6	484	4	US-09-543-681A-5890
13	71	13.5	169	4	US-09-328-352-6931
14	70.5	13.4	188	4	US-09-489-039A-7704
15	70.5	13.4	259	4	US-09-328-352-6237
16	70.5	13.4	514	4	US-09-543-681A-4255
17	70	13.3	342	2	US-08-742-011-2
18	70	13.3	342	3	US-09-275-384B-5
19	70	13.3	342	3	US-09-116-498-2
20	70	13.3	342	3	US-09-116-498-4
21	70	13.3	342	4	US-09-449-437A-2
22	70	13.3	342	4	US-09-517-605-9
23	69.5	13.2	422	4	US-09-489-039A-7295
24	69.5	13.2	424	4	US-09-489-039A-7646
25	69.5	13.2	839	4	US-09-489-039A-13252
26	68.5	13.1	342	3	US-09-116-498-6
27	68.5	13.0	414	4	US-09-328-352-6961

28	67.5	12.8	233	4	US-09-134-000C-6063	Sequence 6063, Ap
29	67	12.7	195	4	US-09-134-000C-5992	Sequence 5992, Ap
30	66.5	12.6	398	4	US-09-134-001C-4353	Sequence 4353, Ap
31	66.5	12.6	421	4	US-09-489-039A-9655	Sequence 9655, Ap
32	66.5	12.6	463	4	US-09-252-991A-2935	Sequence 2935, A
33	66	12.5	269	4	US-09-540-236-2112	Sequence 2112, Ap
34	66	12.5	485	4	US-09-489-039A-9890	Sequence 9890, Ap
35	66	12.5	659	4	US-09-543-681A-6013	Sequence 6013, Ap
36	65.5	12.5	287	4	US-09-252-991A-23091	Sequence 23091, A
37	65.5	12.5	351	4	US-09-134-001C-4387	Sequence 4387, Ap
38	65.5	12.5	440	4	US-09-134-001C-3286	Sequence 3286, Ap
39	65.5	12.5	1053	4	US-09-543-681A-3981	Sequence 5981, Ap
40	65	12.4	333	4	US-09-562-737-10	Sequence 10, Appl
41	65	12.4	404	4	US-09-328-352-6747	Sequence 6747, Ap
42	65	12.4	465	4	US-09-107-532A-5620	Sequence 5620, Ap
43	64.5	12.3	137	4	US-09-489-039A-10320	Sequence 10320, A
44	64.5	12.3	147	4	US-09-732-210-625	Sequence 625, Ap
45	64.5	12.3	309	4	US-09-540-236-3478	Sequence 3478, Ap

ALIGNMENTS

```
RESULT 1
US-09-471-803A-5
Sequence 5, Application US/09471803A
Patent No. 6613545
GENERAL INFORMATION:
APPLICANT: KENNERNECHT, NICOLE
APPLICANT: SAMM, HERMANN
APPLICANT: EGSELM, LOUHA
APPLICANT: PFEFFERLE, WALTER
TITLE OF INVENTION: NUCLEOTIDE SEQUENCE CODING FOR THE EXPORT OF
TITLE OF INVENTION: BRANCHED CHAIN AMINO ACIDS, PROCESS FOR THE
FILE REFERENCE: 21123/265496/MAS
CURRENT APPLICATION NUMBER: US/09/471, 803A
CURRENT FILING DATE: 1999-12-23
PRIOR APPLICATION NUMBER: DE 199 51 708.8
PRIOR FILING DATE: 1999-10-27
NUMBER OF SEQ ID NOS: 12
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 5
LENGTH: 108
TYPE: PRT
ORGANISM: Corynebacterium glutamicum
FEATURE:
OTHER INFORMATION: ATCC14752
US-09-471-803A-5

Query Match      100.0%; Score 526; DB 4; Length 108;
Best Local Similarity 100.0%; Pred. No. 2.8e-58;
Matches 108; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 MTFPSCILLVAVCAVITPRLAAPPILIKPESQFVGMMAMPAGIATITASTFR 60
DB      1 MTFPSCILLVAVCAVITPRLAAPPILIKPESQFVGMMAMPAGIATITASTFR 60
QY      61 SNAIDKLTITFGLTAVAVITVAHLGGRTLSVGAGTIVFGLVNL 108
DB      61 SNAIDKLTITFGLTAVAVITVAHLGGRTLSVGAGTIVFGLVNL 108

RESULT 2
US-09-543-681A-6109
Sequence 6109, Application US/09543681A
Patent No. 6605709
GENERAL INFORMATION:
APPLICANT: GARY BRETON
TITLE OF INVENTION: NOCULEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTUS MIRABII
FILE REFERENCE: 2709.1002-001
CURRENT APPLICATION NUMBER: US/09/543, 681A
```

PD 20-JUN-2001 .
XX
XX 18-DEC-2000; 2000EP-00127688 .
XX
XX 16-DEC-1999; 99JP-00377484 .
PR 07-APR-2000; 2000JP-00159162 .
PR 03-AUG-2000; 2000JP-00280988 .
XX
XX (KYOM) KYOMA HAKKO KOGYO KK .
XX
XX Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;
PI Tateishi N, Senon A, Ikeda M, Ozaki A;
XX
XX WPI; 2001-376931/40 .
DR N-PSDB; AAH67841 .
XX
XX Novel polynucleotides derived from Coryneform bacteria, for identifying
PT mutation point of a gene, measuring expression of a gene, analyzing
PT expression profile or pattern of a gene and identifying homologous gene.
PT
XX
XX Claim 17; SEQ ID NO 6376; 246pp + Sequence Listing; English.
XX
XX The present invention provides a number of nucleotide and protein
CC sequences from the Coryneform bacterium Corynebacterium glutamicum. These
CC are useful for identifying the mutation point of a gene derived from a
CC mutant of coryneform bacterium, measuring expression amount and analysing
CC the expression profile or expression pattern of a gene derived from
CC Coryneform bacterium, and identifying a homologue of a gene derived from
CC Coryneform bacterium. Coryneform bacteria are useful for producing amino
CC acids, nucleic acids, vitamins, saccharides and organic acids.
CC particularly L-lysine. The present sequence is a protein described in the
CC exemplification part of the invention. Note: The sequence data for this patent
CC did not form part of the printed specification, but was obtained in
CC electronic format directly from the European Patent Office
XX
XX Sequence 440 AA;

```
Query Match Similarity      14.5%; Score 76.5; DB 4; Length 440;  
Best Local Similarity     26.1%; Pred. No. 1.1;  
Matches           35; Conservative    19; Mismatches   35; Indels   45; Gaps   7  
  
QY       11 VVAACAVTTFALRAVPFLIKELRESQ-----FVGKKA-MMP----- 47  
          |||::|||::|::|::|::|::|::|::|::|::|::|::|::|:  
DB       202 LVAIVAAWVLFVSAPALLRVPEIEAQVAAEDHPGKLIAAYKDLPGLAEIWKODRNVY 261  
          ::||::|::|::|::|::|::|::|::|::|::|::|::|:  
QY       48 -----AGIL--AIIITASTFRSNALDDLTPLFGLLAVAIVTAHLLGGRRILL 92  
          ||::|::|::|::|::|::|::|::|::|::|::|::|:  
DB       262 FLIAAVERPDGLAGVFTFGAILLAVSYVGHSAADV-LIFGVANVANVASLGALLGG--FL 316  
          ||::|::|::|::|::|::|::|::|::|::|::|::|:  
QY       93 S--VGAGTIVEVGL 104  
          ||::|::|:  
DB       317 DDRVGPKPRIILISL 330
```

Search completed: April 16, 2004, 06:53:15
Job time : 34.3844 secs

CC pathway in which a proliferation-required gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
 CC the target prokaryotic essential genes. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences

XX Sequence 464 AA:
 SQ Query Match 14.6%; Score 77; DB 6; Length 464;
 Best Local Similarity 28.3%; Pred. No. 0.99;
 Matches 30; Conservative 18; Mismatches 38; Indels 20; Gaps 4;

QY 17 VTFPLRAVPFLIK-----PLRE---SQFVKRAMMPAGIILAFSTFR 60
 DB 229 LTFPLSLFAFLLKSHKPMINLFEKIRFTSALVALAOFYIGVITL--PTFF 286
 QY 61 SNAIDKTLTFGLIAVAVTVAHLGGRTLL--SVGATVFVGL 104
 DB 287 TTIGKTELDAAIILFMSIVFICGGLSVINQGLPRLVFGH 332

RESULT 14
 AAB79004
 ID AAB79004 standard; protein; 440 AA.

XX AAB79004;
 AC XX
 DT 30-APR-2001 (first entry)

DE C. glutamicum SRT protein sequence SEQ ID NO:268.

XX fine chemical production; stress; resistance; tolerance; SRT;
 KW nonproteinogenic amino acid; purine base; pyrimidine base; nucleoside;
 KW nucleotide; lipid; saturated fatty acid; unsaturated fatty acid; diol;
 KW carbohydrate; aromatic compound; vitamin; cofactor; polypeptide; enzyme;
 KW evolutionary study; environmental hazard; fermentation.

XX Corynebacterium glutamicum.

XX MO200100804-A2.

XX 04-JAN-2001.

XX 23-JUN-2000; 2000WC-1B000922.

XX 25-JUN-1999; 99US-0141031P.

XX 01-JUL-1999; 99DE-0103042P.

XX 01-JUL-1999; 99US-0142692P.

XX 08-JUL-1999; 99DE-01031413.

XX 08-JUL-1999; 99DE-01031457.

XX 08-JUL-1999; 99DE-01031541.

XX 09-JUL-1999; 99DE-01032209.

XX 14-JUL-1999; 99DE-01032914.

XX 27-AUG-1999; 99DE-01040764.

XX 31-AUG-1999; 99US-0151214P.

XX 99DE-01041382.

XX (BADI) BASF AG.

PI Pompeius M, Kroeger B, Schroeder H, Zelder O, Habenhauer G;
 PI Lee H, Kim H;
 XX WPI; 2001-061972/07.
 DR N-PSDB; AAF71117.
 XX

PT New isolated Corynebacterium glutamicum nucleic acid encoding a stress,
 PT tolerance or resistance protein, for production or modulation of
 PT production of fine chemicals, such as, e.g. amino acids, lipids,
 PT carbohydrates, or enzymes.

XX Claim 20; Page 481-482; 526pp; English.

XX AAF70984 to AAF71133 encode the Corynebacterium glutamicum stress,
 CC resistance, and tolerance (SRT) proteins given in AAB78871 to AAB79020.
 CC The C. glutamicum SRT genes (I) can be used in vectors (II) for
 CC expression in host cells and production of fine chemicals, such as, an
 CC organic acid, a proteinogenic or nonproteinogenic amino acid (preferred),
 CC a purine or pyrimidine base, a nucleoside, a nucleotide, a lipid, a
 CC saturated or unsaturated fatty acid, a diol, a carbohydrate, an aromatic
 CC compound, a vitamin, a cofactor, a polypeptide, or an enzyme. The fine
 CC chemical production can be modulated. The presence of (I) or the SRT
 CC proteins (III) encoded by them are used for diagnosing the presence or
 CC activity of Corynebacterium diptheriae. (I), (II), (III) and host cells
 CC containing them can be used to map the genomes of organisms related to C.
 CC glutamicum, to identify and localise C. glutamicum sequences of interest,
 CC in evolutionary studies, in determination of SRT protein regions required
 CC for function, in modulating the SRT protein activity, and in modulating
 CC the activity of an SRT pathway. (II) are used to permit C. glutamicum to
 CC survive in an environment that is normally environmentally or chemically
 CC hazardous to it. (I) and protein molecules encoded by it increase the
 CC survival of C. glutamicum to chemical and environmental hazards and
 CC provide a means for continued growth and multiplication in large scale
 CC fermentative growth conditions. By increasing the growth rate or
 CC maintaining a normal growth rate in poor or toxic conditions, the yield,
 CC production and/or efficiency or production of fine chemicals from a
 CC culture may be increased

XX Sequence 440 AA;

XX Query Match 14.5%; Score 76.5; DB 4; Length 440;
 XX Best Local Similarity 26.1%; Pred. No. 1.1;
 XX Matches 35; Conservative 19; Mismatches 35; Indels 45; Gaps 7;

QY 11 VVAVCAVTFPALRAVPFLIKPLRESQ-----FVGMA-MMP----- 47

DB 202 LVAVTAAVWFVSAIPALRVPEIEAQAEDHPKGLIAAYKDLFGQIAELMQDRNSVY 261

QY 48 -----AAILT-----AIIITASTFRSNAIDKTLTFGLIAVAVTVAHLGGRTLL 92

DB 262 FLIAATVFRDGLAGVTFGLIALVSYGSLGSDV--LIFVAAVNVGALGALGG---FL 316

QY 93 S--VGATVFVGL 104

DB 317 DDRVGPRIILISL 330

RESULT 15

XX AAG92622 standard; protein; 440 AA.

XX AAG92622;

XX 26-SEP-2001 (first entry)

DE C glutamicum protein fragment SEQ ID NO: 6376.

XX Coryneform bacterium; amino acid synthesis; vitamin; saccharide;

XX organic acid synthesis.

XX Corynebacterium glutamicum.

XX EPI108790-A2.

XX

Query Match 15.0%; Score 79; DB 4; Length 314;
 Best Local Similarity 26.7%; Pred. No. 0.35;
 Matches 20; Conservative 17; Mismatches 26; Indels 12; Gaps 1;
 4 DFCILLVAVCAVITFALRAVFLIKPRLBSQFVGMMAMPAGIATITASTFRSNA 63
 238 DIIDIMVIGIASII-----RPKVDPSIVATPITVLVAILTVSLFRNNK 285
 64 IDLKTITFGLAVAI 78
 286 VGRKTAIVTLLVYSI 300
 Db

RESULT 12

ABM67632
 ID ABM67632 standard; protein; 447 AA.

AC ABM67632;
 XX

DT 20-NOV-2003 (first entry)
 XX

DE Photorhabdus luminescens protein sequence #729.
 XX

XX Antibacterial; fungicide; insecticide; polymorphism; genetic analysis;
 XX detection; food; gene expression; plant; animal; microorganism; toxin;
 XX antibiotic; biopesticide; virulence factor; disease model; plague;
 XX whooping cough.
 XX

OS Photorhabdus luminescens.
 XX

PN WO200294867-A2.
 XX

PD 28-NOV-2002.
 XX

PF 07-FEB-2002; 2002WO-IB003040.
 XX

PR 07-FEB-2001; 2001PR-00001659.
 XX

PA (INSP) INST PASTEUR.
 XX (CNRS) CENT NAT RECH SCI.

PI Duchaud E, Taourit S, Glaser P, Frangeul L, Kunst F, Danchin A;
 XX Buchrieser C;
 XX

DR WPI; 2003-148459/14.
 XX

PT Genomic sequence of Photorhabdus luminescens and encoded polypeptides,
 XX useful e.g. as therapeutic antimicrobials and agricultural pesticides.
 XX

PS Claim 2; SEQ ID NO 729; 1205bp; French.
 XX

XX The invention relates to the isolation of genes and their encoded
 CC proteins from Photorhabdus luminescens. The isolated sequences are
 CC sources of probes and primers for detecting the genome of P. luminescens
 CC and related species; to study polymorphisms; for gene analysis and for
 CC detection/amplification of the genes. Antibodies (Ab) raised against the
 CC polypeptides encoded by the genes are used for detection/identification
 CC of P. luminescens, e.g. in foods. The genes, proteins, Ab and cells that
 CC carry a gene-containing vector are used to select compounds that
 CC modulate, regulate, induce or inhibit expression of the genes in plants,
 CC animals or microorganisms other than P. luminescens and are able to alter
 CC response or sensitivity to toxins and antibiotics produced by P.
 CC luminescens. Cells transformed to express the genes are useful for
 CC recombinant production of the proteins, particularly toxins and
 CC antibacterial agents useful as insecticides, bactericides and fungicides. The
 CC genes, proteins, vectors containing the genes and Ab are also useful
 CC therapeutically (to treat microbial infection by bacteria or fungi that
 CC are sensitive to P. luminescens-encoded toxins or antibiotics) and as
 CC biopesticides. Other uses of the genes and the proteins are as virulence
 CC factors and for identifying targets of human diseases for which P.
 CC luminescens is a model (particularly plague and whooping cough). This
 CC sequence represents one of the isolated P. luminescens proteins

SQ Sequence 447 AA;

Query Match 14.7%; Score 77.5; DB 6; Length 447;
 Best Local Similarity 24.8%; Pred. No. 0.82;
 Matches 34; Conservative 25; Mismatches 47; Indels 31; Gaps 5;

QY 2 TDFSCILLVAVCAVITFALRAVFLIKPRLBSQ-----FVGMMAMPAG 49
 196 STVFGRIIVAVILFLCLISYGLR-IKFTLAGFLRMSQMLKILYIGLPAAGENVVITIQY 254
 50 ILAILRSTFRSNAIDLKTITFGL-----IYAVI---TVVAHLTGR-----RTL 91
 255 MAALAFITGLMGETISLAQTLTYFOLSLFWMLFGIATISIGNEINVGHLVGAKRFEDAYIRGI 314

QY 92 LSVAGTIVFVGLVNLF 108
 315 NSLKIGCIWTIGVTAF 331
 Db

RESULT 13

ABU44538
 ID ABU44538 standard; protein; 464 AA.

AC ABU44538;
 XX

DT 19-JUN-2003 (first entry)
 XX

DE Protein encoded by Prokaryotic essential gene #30065.
 XX

XX Antisense; prokaryotic essential gene; cell proliferation; drug design.
 XX

OS Streptococcus mutans.
 XX

PN WO200277183-A2.
 XX

PD 03-OCT-2002.
 XX

PF 21-MAR-2002; 2002WO-US009107.
 XX

PR 21-MAR-2001; 2001US-00815242.
 XX

PR 06-SEP-2001; 2001US-00946993.
 XX

PR 25-OCT-2001; 2001US-0342923P.
 XX

PR 08-FEB-2002; 2002US-00072851.
 XX

PR 06-MAR-2002; 2002US-0362699P.
 XX

PA (ELIT-) ELITRA PHARM INC.
 XX

PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zykkind JW;
 XX Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
 XX

DR WPI; 2003-029926/02.
 XX

PT New antisense nucleic acids, useful for identifying proteins or screening
 XX for homologous nucleic acids, required for cellular proliferation to
 XX isolate candidate molecules for rational drug discovery programs.
 XX

PS Claim 25; SEQ ID NO 72462; 1766bp; English.
 XX

XX The invention relates to an isolated nucleic acid comprising any one of
 CC the 613 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological

XX Claim 6; SEQ ID NO 665; 192pp; French.
 XX
 CC The present invention relates to the genome sequence of *Listeria*
 CC monocytes EGD-e (see ABA03041). The genome sequence and fragments of
 CC it are useful for selecting probes and primers for detecting genes in *L.*
 CC monocytes and related organisms, and for studying genetic
 CC polymorphisms and other genomes. The present sequence is a protein
 CC encoded by the genome sequence of the present invention. Proteins
 CC expressed from the genome sequence are useful for raising specific
 CC antibodies, identification of *L. monocytes* and related organisms, and
 CC for biosynthesis and biodegradation, especially biosynthesis of Vitamin
 CC B12. The genome sequence and proteins encoded by it are also useful for
 CC selecting compounds that regulate gene expression and cell replication
 CC and modulate *L. monocytes*-related diseases. In addition, the genome
 CC sequence and proteins encoded by it are useful in pharmaceutical and
 CC vaccines compositions for the treatment or prevention of infections by *L.*
 CC monocytes and related organisms. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 CC
 SQ Sequence 108 AA;

Query Match 15.8%; Score 83; DB 5; Length 108;
 Best Local Similarity 26.9%; Pred. No. 0.031;
 Matches 29; Conservative 23; Mismatches 46; Indels 10; Gaps 3;

QY 6 SCILIVAVCAVTFPLRAVPLIKPLRESQFVGKMMNAGLALITASTFRSND 65
 DB 5 SYTLFVIGGLVTFPRVLPFLVRKQLDPVIVYLSVP--LCITLALFVOSLIT 61
 QY 66 LKTLTFG-----LIAVAIVVAHLGSRRTLLSVGAGTIVFV 108
 DB 62 RENSFGINENMLASLPITITAIL--TKNIMIVVVGIIISMLRLF 107

RESULT 10
 ABP40327
 ID ABP40327 standard; protein; 156 AA.
 XX
 AC ABP40327;

DT 24-JUL-2002 (first entry)
 XX
 DE Staphylococcus epidermidis ORF amino acid sequence SEQ ID NO:5172.
 XX
 KM Staphylococcus epidermidis; open reading frame; ORF; bacterial infection;
 XX
 KM antibacterial; gene therapy.

OS Staphylococcus epidermidis.

XX US6380370-B1.

XX 30-APR-2002.

XX 13-AUG-1998; 98US-00134001.

XX 14-AUG-1997; 97US-0055779P.

XX 08-NOV-1997; 97US-0064964P.

XX (GENO-) GENOME THERAPEUTICS CORP.

XX Doucette-Stamm LA, Bush D;

XX WPI; 2002-381255/41.

XX N-PSDB; ABN92872.

XX Novel isolated nucleic acid encoding a *Staphylococcus epidermidis*
 PT polypeptide, useful for diagnosing and treating bacterial infections.

XX Claim 1; SEQ ID NO 5172; 267pp; English.

CC AAB90538 to AAB93374 represent *Staphylococcus epidermidis* open reading
 CC frame (ORF) nucleic acid sequences which encode the amino acid sequences
 CC given in AAB93374 to AAB97960. The *S. epidermidis* sequences have
 CC antibacterial activity and can be used in gene therapy. The sequences can
 CC also be used in the diagnosis and treatment of bacterial infections,
 CC particularly *S. epidermidis* infections. The sequences can be used to
 CC screen for compounds able to interfere with the *S. epidermidis* life cycle
 CC or inhibit *S. epidermidis* infection. N.B. The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from the USPTO web site
 CC
 SQ Sequence 156 AA;

Query Match 15.7%; Score 82.5; DB 5; Length 156;
 Best Local Similarity 25.7%; Pred. No. 0.056;
 Matches 28; Conservative 22; Mismatches 44; Indels 15; Gaps 5;

QY 1 MTDFSCILIVAVCAVTFPLRAVPLIKPLRESQFVGKMMNAGLALITASTFR 60
 DB 48 MTDLHLVLIV-LCGVTLVIVIPFIMISRYNLPAIVIKMISFIP--ITLFTALIT- 102
 QY 61 SNAIDLKTTFG-----LIAVAIVVAHLGSRRTLLSVGAGTIVFV 102
 DB 103 DGVIQGHDAFGYTLNLPYIIALVPTVMAIFRSLTVITLGG--IFV 148

RESULT 11
 AAB96384
 ID AAB96384 standard; protein; 314 AA.

XX AAB96384;

DT 29-OCT-2001 (first entry)

DE Putative *P. abyssi* sodium/calcium antiporter.

XX Hyperthermophilic archaeon; hyperthermophilic protein.

OS *Pyrococcus abyssi*.

XX FR2792651-A1.

XX 27-OCT-2000.

XX 21-APR-1999; 99FR-00005034.

XX 21-APR-1999; 99FR-00005034.

XX (CNRS) CNRS CENT NAT RECH SCI.

XX (IFREMER) IFREMER INST FR RECH EXPL MER.

XX Forterre P, Thierry JC, Prieur D, Dietrich J, Lecompte O;

XX Querellou J, Weissendach J, Saurin W, Heilig R;

XX WPI; 2001-126236/14.

XX New nucleotide sequences isolated from *Pyrococcus abyssi* encode proteins
 PT useful in industry.

XX Claim 7; Page 1068-1069; 1657pp; French.

XX The present invention relates to the genomic sequence of *Pyrococcus*
 CC *abyssi* (see AAF86431 and AAH41223-7) and *P. abyssi* proteins. *P. abyssi* is
 CC a hyperthermophilic archaeon, which is isolated from deep-sea
 CC hydrothermal vents. The present sequence is one such *P. abyssi* protein.
 CC The proteins of the present invention have various potential industrial
 CC uses, since the proteins are stable at very high temperatures, some up to
 CC 110 degrees centigrade. Note: This patent is in the same patent family as
 CC WO200065062, which contains additional sequences as shown in AAB99132-
 CC AAB99143, AAH75903-AAH75920 and AAG66436
 XX
 SQ Sequence 314 AA;

XX SQ Sequence 297 AA;
 Query Match 16.6%; Score 87.5; DB 6; Length 297;
 Best Local Similarity 30.3%; Pred. No. 0.029;
 Matches 37; Conservative 21; Mismatches 39; Indels 25; Gaps 8;
 7 CILVVAVVAVITPALRAVPLILK-----PLRSQFVG---KMMMMRAGILAILTA 56
 140 CPLL---VCLVLTLLMKTEIGLVLTSGDNPMSANGVNDTKIVGMSNGLIALCG 196
 57 SFPRSN-----AIDKLTLPGLIAVAIT-VVAH--LLGGRRTLLSVAGTIVF-VGLVN 106
 197 SLFQNDGSDVTSQGTIVGSLISVILAEVLIHDLITGR--LISIGIAIVRLIILN 254
 QY 107 LF 108
 DB 255 IX 256
 RESULT 6
 ABR42400
 ID ABR42400 standard; protein; 563 AA.
 AC ABR42400;
 XX
 DT 11-AUG-2003 (first entry)
 XX
 DE Human type I adenylylate cyclase splice variant.
 XX
 KW Adenylylate cyclase; enzyme; human; neuroprotective; noctropic;
 KW antiParkinsonian; cerebroprotective; analgesic; antidiabetic;
 XX antiinflammatory; gene therapy.
 XX
 OS Homo sapiens.
 XX
 PN WO2003038083-A1.
 XX
 PD 08-MAY-2003.
 XX
 PF 29-OCT-2002; 2002MO-EP012059.
 XX
 PR 29-OCT-2001; 2001US-0330710P.
 PR 23-JUN-2002; 2002US-0350019P.
 PR 19-JUN-2002; 2002US-0389538P.
 PR 26-SEP-2002; 2002US-0413597P.
 XX
 PA (FARB) BAYER AG.
 XX
 PI Zhu Z;
 XX
 DR WPI; 2003-441359/41.
 DR N-PSDB; ACCS8112.
 XX
 PT New polynucleotide encoding a type I adenylylate cyclase polypeptide,
 PT useful for diagnosing, preventing or treating diseases associated with
 PT type I adenylylate cyclase dysfunction, e.g. CNS disease or diabetes.
 XX
 PS Claim 1; Page 139-140; 167pp; English.
 XX
 CC The present sequence is the protein sequence of a splice variant of human
 CC type I adenylylate cyclase (see also ABR42399), an enzyme that catalyzes
 CC the synthesis of cAMP from ATP. The invention provides reagents that
 CC regulate human type I adenylylate cyclase and reagents which bind human
 CC type I adenylylate cyclase gene products. These reagents, and vectors
 CC containing polynucleotides encoding e.g. present sequence, can be used to
 CC modulate the activity of human type I adenylylate cyclase in a disease,
 CC especially a central nervous system (CNS) disorder, diabetes and chronic
 CC obstructive pulmonary disorder (claimed). The CNS disorders may include
 CC Alzheimer's disease, Parkinson's disease, stroke or neuropathic pain.
 CC Type I adenylylate cyclase can be obtained by recombinant methods, and used
 CC to screen for activator and inhibitor compounds

XX SQ Sequence 563 AA;
 Query Match 16.2%; Score 85; DB 6; Length 563;
 Best Local Similarity 31.4%; Pred. No. 0.13;
 Matches 33; Conservative 18; Mismatches 44; Indels 10; Gaps 5;
 12 VAVCAVITPALRAVPLILKPLRSQFVGKMMMPAGILAILTASTPRSNAI---DK 67
 125 LALLFSLTFLALCCPPALGGPARGSAGAGPATAEQGWQOLLV-TFVSVALLPYRSIL 183
 68 TLTFGLIAVA--ITVAHLLGGRRTLL--SVAGTIVFVGVNLF 108
 184 AIGGLVAASHLLVTATLVPAAPRLMRTLGANLLFVG-VNMY 227
 QY
 DB
 RESULT 7
 AAE17133
 ID AAE17133 standard; protein; 1119 AA.
 AC AAE17133;
 XX
 DT 22-APR-2002 (first entry)
 XX
 DE Human adenylyl and guanylyl cyclase (ADGUC)-5.
 XX
 KW Human; adenylyl and guanylyl cyclase; ADGUC-5; cardiovascular disorder;
 KW angina pectoris; myocardial infarction; vision disorder; keratitis;
 KW iritis; cataract; neurological disorder; epilepsy; Alzheimer's disease;
 KW Pick's disease; stroke; mental disorder; mood and anxiety disorder;
 KW reproductive disorder; infertility; endometriosis; impotence; asthma;
 KW smooth muscle disorder; migraine; bacterial infection; gene therapy;
 KW transgenic animal; vaccine; enzyme.
 XX
 OS Homo sapiens.
 XX
 PN WO200202757-A2.
 XX
 PD 10-JAN-2002.
 XX
 PF 26-JUN-2001; 2001MO-US020491.
 XX
 PR 29-JUN-2000; 2000US-0215476P.
 PR 04-AUG-2000; 2000US-0223545P.
 PR 31-AUG-2000; 2000US-0229876P.
 PR 22-SEP-2000; 2000US-0234838P.
 PR 29-SEP-2000; 2000US-0236483P.
 XX
 PA (INCY-) INCYTE GENOMICS INC.
 XX
 PI Gandhi AR, Tribouley C, Ding L, Lu DM, Lee EA, Yue H, Yang J,
 PI Baughn MR, Thornton M, Yao MG, Walla NK, Tang YT, Elliott VS, Lu Y;
 DR WPI; 2002-154740/20.
 DR N-PSDB; AMD28061.
 XX
 PT Novel human adenylyl and guanylyl cyclases and polynucleotides encoding
 PT the cyclases, useful for treating, diagnosing or preventing
 PT cardiovascular, neurological, vision, reproduction and smooth muscle
 PT disorders.

```

Query March 24.7%; Score 130; DB 4; Length 115;
Best Local Similarity 35.0%; Pred. No. 5.1e-08;
Matches 36; Conservative 16; Mismatches 47; Indels 4; Gaps 2;

QY      8 ILTVAV---CAVTFALRAVFFLLKPLRESQFVGKMMMPAGIIAIIATSTFRSNAI 64
DB      12 ILTVAAVLIPAIITILLIRIPFFAMKRVNSNQMLGVLGRTPVGVMVIVLYTLFGQVS 71

QY      65 DLKTIITFGLIAVITVAHLLGGRRTLLSVAGTIVFVGLNL 107
DB      72 APGGVGSALIAVAFALHMLKXGASGLIVG-GTLAYMFLVNV 113

RESULT 4
ID      ABB54958 standard; protein; 108 AA.
AC      ABB54958;
AD      29-AUG-2003 (revised)
DT      16-MAY-2002 (first entry)
XX      Lactococcus lactis protein yqfC.
DE      Lactococcus lactis protein yqfC.
XX      Biosynthesis; biodegradation; lactic bacterium; yogurt; cheese.
XX      Lactococcus lactis; IL1403.
OS      Lactococcus lactis; IL1403.
XX      FR2807446-A1.
XX      12-OCT-2001.
PD      11-APR-2000; 2000FR-00004630.
XX      11-APR-2000; 2000FR-00004630.
XX      11-APR-2000; 2000FR-00004630.
PR      (INRG ) INRA INST NAT RECH AGRONOMIQUE.
PA      Bolotine A, Sorokine A, Renault P, Ehrlich SD;
XX      WPI; 2002-043418/06.
XX      New nucleotide sequence useful in the identification of Lactococcus
PT      lactis and related species.
XX      Claim 6; SEQ ID NO 1660; 2504bp; French.
XX      The present invention is related to a Lactococcus lactis nucleotide
CC      sequence (ABA90521) and related proteins (ABB53300-ABB55621). The nucleic
CC      acid sequence is useful in the detection and/or amplification of related
CC      acid sequence, particularly to identify Lactococcus lactis or related
CC      species. The proteins of the invention are useful for the biosynthesis or
CC      biodegradation of a composition of interest. The invention helps research
CC      in lactic bacteria, particularly useful in the production of yogurt and
CC      cheese. Note: The sequence data for this patent is based on equivalent
CC      patent WO200117334 (published 18-OCT-2001) which is available in
CC      electronic format directly from WIPO at
CC      ftp.wipo.int/pub/published_pct_sequences. (Updated on 29-AUG-2003 to
CC      standardise OS field)
XX      Sequence 108 AA;

Query Match 18.0%; Score 94.5; DB 5; Length 108;
Best Local Similarity 29.9%; Pred. No. 0.0012;
Matches 32; Conservative 24; Mismatches 42; Indels 9; Gaps 4

QY      3 TDSCILLVAVCAVTFALRAVFFLLKPLRESQFVGKMMMPAGIIAIIATST-FRS 61.
DB      2 SSPEFISLITIGCAITWISRVLPFILKKMSIPQIVVEYLSFVPVIMSAIMISNLFIO 61

QY      62 NALDILTLTF-GLIAVITVAHLLGGRRTLLSVAGTIVFVGLNL 107
DB      62 HIGLPSVMMNNILASIPVLAAIL--TKNUL-----VILVGVFSL 101

```

RESULT 5
 ABU45714
 ID ABU45714 standard; protein: 297 AA.
 XX
 XX ABU45714;
 XX
 DT 19-JUN-2003 (first entry)
 DE Protein encoded by Prokaryotic essential gene #31241.
 XX
 XX Antisense; prokaryotic essential gene; cell proliferation; drug design.
 XX
 OS Streptococcus pneumoniae.
 XX
 PN WO200277183-A2.
 XX
 PD 03-OCT-2002.
 XX
 PE 21-MAR-2002; 2002WO-US009107.
 XX
 PR 21-MAR-2001; 2001US-00815242.
 XX
 PR 06-SEP-2001; 2001US-00948993.
 PR 25-OCT-2001; 2001US-0342923P.
 PR 08-FEB-2002; 2002US-00072851.
 PR 06-MAR-2002; 2002US-0352699P.
 XX
 PA (ELIT-) ELITRA PHARM INC.
 XX
 PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
 PI Wall D, Trawick JD, Carr GC, Yamamoto R, Forsyth RA, Xu HH;
 XX
 DR WPI; 2003-029926/02.
 XX
 DR N-PSDB; ACA9584.
 XX
 PT New antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.
 XX
 PS Claim 25; SEQ ID NO 73638; 1766P; English.
 XX
 XX The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying the target of a biological pathway
 CC the gene product of that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC or a gene on which the test compound directly inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
 CC the target prokaryotic essential genes. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pat_sequences

CC especially overexpressed; and (d) method for isolating the brnE and/or
CC brnF genes. (1) is used for transformation of coryneform bacteria being
CC used for fermentative production of branched-chain amino acids,
CC specifically leucine, isoleucine and valine, which are useful in medicine
CC and animal nutrition. (1) can also be used as source of primers and
CC probes for isolation of related sequences. Transformation with (1)
CC increases yield of branched-chain amino acids. This sequence represents
CC the Corynebacterium glutamicum ATCC 14752 brnE protein described in the
CC method of the invention
CC
XX

Sequence 108 AA;

Query Match 100.0%; Score 526; DB 4; Length 108;

Best Local Similarity 100.0%; Pred. No. 4.1e-57; Mismatches 0; Indels 0; Gaps 0;

Matches 108; Conservative 0;

QY 1 MTTDFSCILLVAVCAVITFALRAVPFLIKPFRSQFVGKAMMPAGILAITASTFR 60
1 MTTDFSCILLVAVCAVITFALRAVPFLIKPFRSQFVGKAMMPAGILAITASTFR 60
DB 61 SNAIDKTLTFGLIAVAITVAHLIGRRITLISVGAGTIVFGVNLNF 108

QY 61 SNAIDKTLTFGLIAVAITVAHLIGRRITLISVGAGTIVFGVNLNF 108
61 SNAIDKTLTFGLIAVAITVAHLIGRRITLISVGAGTIVFGVNLNF 108
DB

RESULT 2

AAG90036 ID AAG90036 standard; protein; 108 AA.

XX AAG90036;

DT 26-SEP-2001 (first entry)

DE C glutamicum protein fragment SEQ ID NO: 3790.

XX Coryneform bacterium; amino acid synthesis; vitamin; saccharide;

XX organic acid synthesis.

OS Corynebacterium glutamicum.

XX EPI108790-A2.

XX 20-JUN-2001.

PF 18-DEC-2000; 2000EP-00127688.

XX 16-DEC-1999; 99JP-00377484.

PR 07-APR-2000; 2000JP-00159162.

PR 03-AUG-2000; 2000JP-00280988.

XX (KYOW) KYOWA HAKKO KOGYO KK.

XX Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;

PI Tateishi N, Senoh A, Ikeda M, Ozaki A;

PT WPI; 2001-376931/40.

DR N-PSDB; AAH65255.

XX Novel polynucleotides derived from Coryneform bacteria, for identifying

PT mutation point of a gene, measuring expression of a gene, analyzing

XX expression profile or pattern of a gene and identifying homologous gene.

XX Claim 17; SEQ ID NO 3790; 246bp + Sequence Listing; English.

XX The present invention provides a number of nucleotide and protein

CC sequences from the Coryneform bacterium Corynebacterium glutamicum. These

CC are useful for identifying the mutation point of a gene derived from a

CC mutant of coryneform bacterium, measuring expression amount and analyzing

CC the expression profile or expression pattern of a gene derived from

CC Coryneform bacterium, and identifying a homologue of a gene derived from

CC Coryneform bacterium. Coryneform bacteria are useful for producing amino

CC acids, nucleic acids, vitamins, saccharides and organic acids,

CC particularly L-lysine. The present sequence is a protein described in the

CC exemplification of the invention. Note: The sequence data for this patent

CC did not form part of the printed specification, but was obtained in

CC electronic format directly from the European Patent Office

XX Sequence 108 AA;

Query Match 100.0%; Score 526; DB 4; Length 108;

Best Local Similarity 100.0%; Pred. No. 4.1e-57; Mismatches 0; Indels 0; Gaps 0;

Matches 108; Conservative 0;

QY 1 MTTDFSCILLVAVCAVITFALRAVPFLIKPFRSQFVGKAMMPAGILAITASTFR 60
1 MTTDFSCILLVAVCAVITFALRAVPFLIKPFRSQFVGKAMMPAGILAITASTFR 60
DB 61 SNAIDKTLTFGLIAVAITVAHLIGRRITLISVGAGTIVFGVNLNF 108

QY 61 SNAIDKTLTFGLIAVAITVAHLIGRRITLISVGAGTIVFGVNLNF 108
61 SNAIDKTLTFGLIAVAITVAHLIGRRITLISVGAGTIVFGVNLNF 108
DB

RESULT 3
AAG93153 ID AAG93153 standard; protein; 115 AA.

XX AAG93153;

DT 26-SEP-2001 (first entry)

DE C glutamicum protein fragment SEQ ID NO: 6907.

XX Coryneform bacterium; amino acid synthesis; vitamin; saccharide;

XX organic acid synthesis.

OS Corynebacterium glutamicum.

XX EPI108790-A2.

XX 20-JUN-2001.

PF 18-DEC-2000; 2000EP-00127688.

XX 16-DEC-1999; 99JP-00377484.

PR 07-APR-2000; 2000JP-00159162.

PR 03-AUG-2000; 2000JP-00280988.

XX (KYOW) KYOWA HAKKO KOGYO KK.

XX Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;

PI Tateishi N, Senoh A, Ikeda M, Ozaki A;

PT WPI; 2001-376931/40.

DR N-PSDB; AAH68372.

XX Novel polynucleotides derived from Coryneform bacteria, for identifying

PT mutation point of a gene, measuring expression of a gene, analyzing

XX expression profile or pattern of a gene and identifying homologous gene.

XX Claim 17; SEQ ID NO 6907; 246bp + Sequence Listing; English.

XX The present invention provides a number of nucleotide and protein

CC sequences from the Coryneform bacterium Corynebacterium glutamicum. These

CC are useful for identifying the mutation point of a gene derived from a

CC mutant of coryneform bacterium, measuring expression amount and analyzing

CC the expression profile or expression pattern of a gene derived from

CC Coryneform bacterium, and identifying a homologue of a gene derived from

CC Coryneform bacterium. Coryneform bacteria are useful for producing amino

CC acids, nucleic acids, vitamins, saccharides and organic acids,

CC particularly L-lysine. The present sequence is a protein described in the

CC exemplification of the invention. Note: The sequence data for this patent

CC did not form part of the printed specification, but was obtained in

CC electronic format directly from the European Patent Office

XX Sequence 115 AA;

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 16, 2004, 06:51:12 ; Search time 30.3844 Seconds
(without alignments)
1004.303 Million cell updates/sec

Title: US-10-608-504-5

Perfect score: 526
Sequence: 1 MTTTPSCILVAVCAVITF.....RTLSVGAGTIVFGVNLNF 108

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*
1: geneseqp19808:*
2: geneseqp19908:*
3: geneseqp20008:*
4: geneseqp20018:*
5: geneseqp20028:*
6: geneseqp20038:*
7: geneseqp20048:*
8: geneseqp20058:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	526	100.0	108	4	AA86248 C. glutam
2	526	100.0	108	4	AA86248 C. glutam
3	130	24.7	115	4	AA86248 C. glutam
4	94.5	18.0	108	5	AB854958 Lactococ
5	87.5	16.6	297	6	ABU45714 Protei
6	85	16.2	563	6	ABR42400 Human
7	85	16.2	1119	5	AAE71333 Human
8	85	16.2	1119	5	ABR42399 Human
9	83	15.8	108	5	ABR42399 Human
10	82.5	15.7	156	5	ABP40327 Staphy
11	79	15.0	314	4	AA86384 Putativ
12	77.5	14.7	447	6	ABM67632 Photoc
13	77	14.6	464	6	ABU44538 Protei
14	76.5	14.5	440	4	AA879004 C. glut
15	76.5	14.5	440	4	AA879004 C. glut
16	75	14.3	415	6	ABU71311 Protein
17	75	14.3	1134	7	AA862005 Adenyly
18	75	14.3	1134	7	AA862005 Adenyly
19	74.5	14.2	379	1	AA870461 Sequence
20	74	14.1	207	4	AB859858 S. cinna
21	74	14.1	227	6	ABP77220 N. gonorr
22	74	14.1	254	4	AA871711 Human
23	74	14.1	312	5	ABP51594 Human
24	74	14.1	312	5	ABP51594 Human
25	74	14.1	313	6	ABU11194 Human

26	74	14.1	313	7	ADC85887	ADC85887 Human GPC
27	74	14.1	317	6	ABR01631	ABR01631 Human G P
28	74	14.1	318	4	AA872535	AA872535 Human OR
29	74	14.1	318	4	AA872225	AA872225 Human OLF
30	74	14.1	330	7	ADC86129	ADC86129 Human GPC
31	73.5	14.0	187	6	ADB12425	ADB12425 A1101ococ
32	73.5	14.0	443	6	ABU24119	ABU24119 Protein e
33	73.5	14.0	1049	2	AA832855	AA832855 E. coli a
34	73.5	14.0	1049	2	AAU34479	AAU34479 E. coli c
35	73.5	14.0	1049	6	ABU15219	ABU15219 Protein e
36	73	13.9	109	6	ABM72816	ABM72816 Staphyloc
37	72.5	13.8	326	7	ADC94038	ADC94038 E. faecit
38	72.5	13.8	1048	6	ABU32055	ABU32055 Protein e
39	72.5	13.8	1048	6	ABU28297	ABU28297 Protein e
40	72	13.7	117	7	ADC95966	ADC95966 E. faecit
41	72	13.7	464	4	ABR67164	ABR67164 Drosophi
42	72	13.7	558	6	ABR42402	ABR42402 Human
43	72	13.7	704	6	ABU21547	ABU21547 Protein e
44	72	13.7	1114	6	ABR42401	ABR42401 Human
45	71.5	13.6	1049	4	AAU38126	AAU38126 Salmonel

ALIGNMENTS

RESULT 1
ID AAB86248 standard; protein; 108 AA.
XX AAB86248;
AC
XX 05-SEP-2001 (first entry)
DT
XX
DE C. glutamicum brnE protein.
XX
XX L-amino acid production; brnF; brnG; branched-chain amino acid;
XX L-amino acid production; leucine; isoleucine; valine; medicine;
XX animal nutrition.
XX
XX Corynebacterium glutamicum.
XX
XX EPI096010-A1.
XX
XX 02-MAY-2001.
XX
XX 11-OCT-2000; 2000EP-00122057.
XX
XX 27-OCT-1999; 99DE-01051708.
XX
XX (DEGS) DEGUSSA AG.
XX (KERJ) FORSCHUNGSZENTRUM JUELICH GMBH.
XX
XX Kennertknecht N, Eggeling L, Sahn H, Pfeifferle W,
XX WPI; 2001-391595/42.
XX DR N-PSDB; AAH21109, AAH21111.
XX
XX New export genes from coryneform bacteria, useful for increasing
XX Fermentative production of branched-chain amino acids.
XX
XX Claim 1a; Page 17; 23pp; German.
XX
XX This invention describes a novel isolated polynucleotide (I) containing
XX at least one sequence that (i) is 70% identical with a sequence that
XX encodes at least one of 251 amino acid (aa) (3) or 108 aa (5)
XX or (5); (iii) is the complement of (i) or (ii), or (iv) contains at least
XX 15 consecutive bases from (i)-(iii). The invention also describes (a)
XX protein derived from sequences of 753 bp (2) or 324 bp (4); (b)
XX coryneform microorganisms, especially Corynebacterium, transformed with
XX one or more (II), where these are replicative DNA; (c) production of
XX branched-chain L-aa by fermentation of coryneform bacteria in which the
XX brnE and/or brnF genes (or equivalent sequences) are amplified.

```

Db      555 GATATATACCAAGAGTTTTCATTGCTTGGAAATTCACATTATGAGTTTTCAG 614
Qy      194 ATCTGAAGACTCTAACCTTTGGCTTCATTCACCTGGTTCAGATTACAGTGGCGCATCTTC 253
Db      615 CTCAGAGCAAGATTAATGATCATGCTGTATATAGCCGGGTTGCATTGTTTGCACTT 674
Qy      254 TTGGCGGTGACGACCACTTGTGAGCGTTGGCGCTGG 290
Db      675 TTTTACTTGCTTTTGATGCTGAAGCTAGGCTGGGG 711

```

RESULT 15

```

CK295424      936 bp      mRNA      linear      EST 15-DEC-2003
LOCUS      EST758138 Nicotiana benthamiana mixed tissue cDNA library,
DEFINITION      normalized, full-length Nicotiana benthamiana cDNA clone NEMCU77 5'
end, mRNA sequence.

```

```

ACCESSION      CK295424
VERSION      CK295424.1 GI:39879801
KEYWORDS      EST.
SOURCE      Nicotiana benthamiana
ORGANISM      Nicotiana benthamiana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamiales; Solanales; Solanaceae; Nicotiana.

```

```

REFERENCE      Buell, C.R., Hart, A., Zismann, V., Karayancheva, S.A., Day, B.,
AUTHORS      1 (bases 1 to 936)

```

```

TITLE      Generation of EST sequences from Nicotiana benthamiana
JOURNAL      Unpublished (2003)
COMMENT      Other ESTs: EST738139

```

```

The Institute for Genomic Research
9712 Medical Center Dr, Rockville, MD 20850, USA
Email: potato-aray@tigr.org
Clones can be requested from TIGR via potato@tigr.org
Seq primer: ATT TNG GTG ACA CTA TAG.
Location/Qualifiers

```

```

FEATURES
source
1..936
/organism="Nicotiana benthamiana"
/mol_type="mRNA"
/db_xref="taxon:4100"
/clone="NEMCU77"
/tissue_type="abiotic and biotic stress-treated leaves,
callus tissue and root tissue"
/lab_host="DH10B-TONA"
/lab_lib="Nicotiana benthamiana mixed tissue cDNA
library, normalized, full-length"
/notes="Vector: PCW/SPORT6.1; Site 1: EcoRI; Site 2: NotI;
supplier: RNA was isolated from Nicotiana benthamiana
tissues that include callus, roots from liquid culture
grown plants, heat-stressed leaves (38 C, 3 hr and 6 hr),
cold-stressed leaves (5 C 3 hr, 6hr), and pathogen
challenged leaves (Pseudomonas syringae pv tomato 12 hr;
Xanthomonas campestris pv campestris 12 hr, 18hr;
Pseudomonas syringae pv phaseolicola 18hr, and Xanthomonas
campestris pv vesicatoria 18hr). RNA was isolated from
these tissues and pooled in approximately equal molar
amounts."

```

ORIGIN

```

Query Match      10.7%; Score 34.6; DB 14; Length 936;
Best Local Similarity 47.5%; Pred. No. 26;
Matches 103; Conservative 0; Mismatches 114; Indels 0; Gaps 0;

```

```

Qy      74 TTGCGTTCTTAATCCCTAGCCCTAGTAATCACATTTGGGCAAAATGCGATGT 133
Db      419 TTCTAATGTTAATGGTCAGCCGAAACATCTCAAGTGGCGGGAATTTGCCCTGT 478
Qy      134 GGATGCCAGCAGGAATCTTGCAATTTTGAACCGCATCAAGTTTGCAGCATGCGATAG 193
Db      479 GGATGATACCAAGTGTTCATTGCTTTGGAATTCCAATTATGAAGTTTTCAGAG 538

```

```

Qy      194 ATCTGAAGACTCTAACCTTTGGCTTCATTCACCTGGTTCAGATTACAGTGGCGCATCTTC 253
Db      539 CTCAGAGCAAGATTAATGATCATGCTGTATATAGCCGGGTTGCATTGTTTGCACTT 598
Qy      254 TTGGCGGTGACGACCACTTGTGAGCGTTGGCGCTGG 290
Db      599 TTTTACTTGCTTTTGATGCTGAAGCTAGGCTGGGG 635

```

```

Search completed: April 17, 2004, 20:01:47
Job time : 1053.18 secs

```

QY 74 TTCGTTCTTAATCCTTAAGCCCTACGTAATACAAATTTGGGCAAAATGGCATGT 133
 DB 404 TTCTAATGTATATCGGTGACGCCGAAAACATCTCAAGTGGCGGGAATTCGCGCTGT 463
 QY 134 GGATGCCAGCAGAAATCCTTGGCATTTGACCGCATCAAGTTTGGAGCAATGGCATAG 193
 DB 464 GGATGATACCAAGATGTTTCAATTTGTTGGAATTCCTCAATTAAGATTTTGCAG 523
 QY 194 ATCTGAAGACTTAACCTTTGATCTCATTTGCCGTTGCGATTACAGTGGCGCATCTTC 253
 DB 524 CTCAGAGCAAGATATAGTCAATGCGTGTATTAGCGGGGTTGCAATGTTTGCATCTT 583
 QY 254 TTGGCGGTGACAGCACTTTGTTGAGCGTTGGCGCTGG 290
 DB 584 TTTTACTTGGCTTTTGATGCTGAAGCTAGGCTGGGG 620

RESULT 13

LOCUS CK288648 890 bp mRNA linear EST 15-DEC-2003
 DEFINITION Nicotiana benthamiana mixed tissue cDNA library, normalized, full-length Nicotiana benthamiana cDNA clone NBMX102.5', end, mRNA sequence.

ACCESSION CK288648
 VERSION CK288648.1 GI:39866370

KEYWORDS
 SOURCE Nicotiana benthamiana
 ORGANISM Nicotiana benthamiana

REFERENCE
 AUTHORS Buell, C.R., Hart, A., Zismann, V., Karameycheva, S.A., Day, B., Staekawicz, B., Jin, H., and Baker, B.
 TITLE Generation of EST sequences from Nicotiana benthamiana Unpublished (2003)
 JOURNAL Contact: Robin Buell
 COMMENT The Institute for Genomic Research 9712 Medical Center Dr, Rockville, MD 20850, USA
 Email: potato-array@tigr.org
 Clones can be requested from TIGR via potato@tigr.org
 Seq primer: ATT TAG GTG ACA CTA TAG.

FEATURES

source
 1..890
 /organism="Nicotiana benthamiana"
 /mol_type="mRNA"
 /db_xref="taxon:4100"
 /clone="NBMX102"
 /tissue_type="abiotic and biotic stress-treated leaves, callus tissue and root tissue"
 /lab_host="DH10B-TonA"
 /clone_lib="Nicotiana benthamiana mixed tissue cDNA library, normalized, full-length"
 /note="Vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI; supplier: RNA was isolated from Nicotiana benthamiana tissues that include callus, roots from liquid culture grown plants, heat-stressed leaves (38 C, 3 hr and 6 hr), cold-stressed leaves (5 C 3 hr, 6hr), and pathogen challenged leaves (Pseudomonas syringae pv tomato 12 hr; Xanthomonas campestris pv campestris 12 hr, 18hr; Pseudomonas syringae pv phaseolicola 18hr, and Xanthomonas campestris pv vesicatoria 18hr). RNA was isolated from these tissues and pooled in approximately equal molar amounts."

ORIGIN

Query Match 10.7%; Score 34.6; DB 14; Length 890;
 Best Local Similarity 47.5%; Pred. No. 25;
 Matches 103; Conservative 0; Mismatches 114; Indels 0; Gaps 0;

QY 74 TTCGTTCTTAATCCTTAAGCCCTACGTAATACAAATTTGGGCAAAATGGCATGT 133
 DB 543 TTCTAATGTATATCGGTGACGCCGAAAACATCTCAAGTGGCGGGAATTCGCGCTGT 602

QY 134 GGATGCCAGCAGAAATCCTTGGCATTTGACCGCATCAAGTTTGGAGCAATGGCATAG 193
 DB 603 GGATGATACCAAGATGTTTCAATTTGTTGGAATTCCTCAATTAAGATTTTGCAG 662
 QY 194 ATCTGAAGACTTAACCTTTGATCTCATTTGCCGTTGCGATTACAGTGGCGCATCTTC 253
 DB 663 CTCAGAGCAAGATATAGTCAATGCGTGTATTAGCGGGGTTGCAATGTTTGCATCTT 722
 QY 254 TTGGCGGTGACAGCACTTTGTTGAGCGTTGGCGCTGG 290
 DB 723 TTTTACTTGGCTTTTGATGCTGAAGCTAGGCTGGGG 759

RESULT 14

LOCUS CK295854 922 bp mRNA linear EST 15-DEC-2003
 DEFINITION Nicotiana benthamiana mixed tissue cDNA library, normalized, full-length Nicotiana benthamiana cDNA clone NBMX72.5', end, mRNA sequence.

ACCESSION CK295854
 VERSION CK295854.1 GI:3980657

KEYWORDS
 SOURCE Nicotiana benthamiana
 ORGANISM Nicotiana benthamiana

REFERENCE
 AUTHORS Buell, C.R., Hart, A., Zismann, V., Karameycheva, S.A., Day, B., Staekawicz, B., Jin, H., and Baker, B.
 TITLE Generation of EST sequences from Nicotiana benthamiana Unpublished (2003)
 JOURNAL Other ESTs: EST758569
 COMMENT Contact: Robin Buell
 The Institute for Genomic Research 9712 Medical Center Dr, Rockville, MD 20850, USA
 Email: potato-array@tigr.org
 Clones can be requested from TIGR via potato@tigr.org
 Seq primer: ATT TAG GTG ACA CTA TAG.

FEATURES

source
 1..922
 /organism="Nicotiana benthamiana"
 /mol_type="mRNA"
 /db_xref="taxon:4100"
 /clone="NBMX72"
 /tissue_type="abiotic and biotic stress-treated leaves, callus tissue and root tissue"
 /lab_host="DH10B-TonA"
 /clone_lib="Nicotiana benthamiana mixed tissue cDNA library, normalized, full-length"
 /note="Vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI; supplier: RNA was isolated from Nicotiana benthamiana tissues that include callus, roots from liquid culture grown plants, heat-stressed leaves (38 C, 3 hr and 6 hr), cold-stressed leaves (5 C 3 hr, 6hr), and pathogen challenged leaves (Pseudomonas syringae pv tomato 12 hr; Xanthomonas campestris pv campestris 12 hr, 18hr; Pseudomonas syringae pv phaseolicola 18hr, and Xanthomonas campestris pv vesicatoria 18hr). RNA was isolated from these tissues and pooled in approximately equal molar amounts."

ORIGIN

Query Match 10.7%; Score 34.6; DB 14; Length 922;
 Best Local Similarity 47.5%; Pred. No. 26;
 Matches 103; Conservative 0; Mismatches 114; Indels 0; Gaps 0;

QY 74 TTCGTTCTTAATCCTTAAGCCCTACGTAATACAAATTTGGGCAAAATGGCATGT 133
 DB 495 TTCTAATGTATATCGGTGACGCCGAAAACATCTCAAGTGGCGGGAATTCGCGCTGT 554
 QY 134 GGATGCCAGCAGAAATCCTTGGCATTTGACCGCATCAAGTTTGGAGCAATGGCATAG 193

ORIGIN

/clone="CSODI052YB18"
 /tissue_type="PLACENTA COT 25-NORMALIZED"
 /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
 /note="1st strand cDNA was primed with a NotI-oligo (dT)
 primer. Five prime end enriched, double-strand cDNA was
 digested with Not I and cloned into the Not I and EcoR V
 sites of the pCMVSPORT 6 vector. Library was normalized."

Query Match 10.7%; Score 34.8; DB 13; Length 1201;
 Best Local Similarity 39.9%; Pred. No. 25; Indels 0; Gaps 0;
 Matches 63; Conservative 25; Mismatches 70; Gaps 0;

QY 154 GCCATTTGACCGCATCAACGTTTGCAGCAATGCGATAGATCTAGACCTTAACCTTT 213
 DB 943 GACTTTCTWCGAGCCGCGCGCGTGGTGTGTCGGGAGCGCGCCCTTCCCGAGT 1002
 QY 214 GGTCTCATTCGCGTTCGATACAGTGTGGCGCATCTTCTGGGGTGCAGCGACCTTG 273
 DB 1003 GATKSGGTTSMGCGCCCTTTKAMAGTGCAGGAGTATTTTCTTCCCGCCGCTC 1062
 QY 274 TTGAGCGTTGGCGCTGGACCAATCGTTTGTGTGACT 311
 DB 1063 CAGKTAGTGGGGGGGSGTACVATCTTCGTSTTAAAGT 1100

RESULT 11 413 bp mRNA linear EST 30-APR-2003
 LOCUS CB239541
 DEFINITION RSH3B11 two-month-old roots from clone 'Beaupre' grown for 19 days
 under restricted irrigation Populus balsamifera subsp. trichocarpa
 x Populus deltoides cDNA 5', mRNA sequence.

ACCESSION CB239541 GI:30233030
 VERSION CB239541
 KEYWORDS
 SOURCE
 ORGANISM

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT
 tags
 Unpublished (2003)
 Contact: Martin FM
 Equipe de Microbiologie Forestiere
 Institut National de la Recherche Agronomique
 Centre INRA de Nancy, 54280 Champenoux, France
 Tel: +33 383 39 40 80
 Fax: +33 383 39 40 69
 Email: fmarin@nancy.inra.fr
 Insert length: 413 Std Error: 0.00
 Seq primer: Fornat 5' AAGCGGCGCATTCGTGTGTGATACC.
 Location/Qualifiers

FEATURES

source

1. .413
 /organism="Populus balsamifera subsp. trichocarpa x
 Populus deltoides"
 /mol_type="mRNA"
 /cultiVar="Beaupre"
 /db_xref="taxon:3695"
 /dev_stage="two-month-old"
 /clone_lib="two-month-old roots from clone 'Beaupre' grown
 for 19 days under restricted irrigation"
 /note="Organ: root; Vector: pTIP1EX2; cDNA library of
 roots from two-month-old Populus trichocarpa Torr. & Gray
 x deltoides Bartr. Ex Marshall (clone 'Beaupre') grown for
 19 days under restricted irrigation to reach 50% of the
 transpiration rate of fully watered plants. The cDNA
 library was constructed from 1 ug of total RNA using the
 SMART cDNA synthesis kit (Clontech, Palo Alto, CA, USA)
 according to the manufacturer's instructions. The
 resulting cDNA was packed into lambda phages using the

ORIGIN

Gigapack III Gold packaging kit (Stratagene, La Jolla,
 CA). The pTIP1EX2 phagemid clones in Escherichia coli
 were obtained by using the mass in vivo excision protocol
 according to the manufacturer's instructions (Clontech)."

Query Match 10.7%; Score 34.6; DB 14; Length 413;
 Best Local Similarity 60.4%; Pred. No. 19;
 Matches 55; Conservative 0; Mismatches 36; Indels 0; Gaps 0;

QY 65 TCCGGGCGGTTCCGTTCTTAATCTTACGCCCTACGTAATACAAATTGTGGGCAAA 124
 DB 71 TCCAGAGCTCCCTCTGAGAGCTATTACCCCTCCAGATTACCTTCAAGTCAAA 130
 QY 125 TGGCGATGTGATGCGACGAGATCTTGC 155
 DB 131 TGGGGTTCGCGCGGCGCACCGAGTTCCTGAC 161

RESULT 12 853 bp mRNA linear EST 15-DEC-2003
 LOCUS CK297477
 DEFINITION EST760191 Nicotiana benthamiana mixed tissue cDNA library,
 normalized, full-length Nicotiana benthamiana cDNA clone NBMD912 5'
 end, mRNA sequence.

ACCESSION CK297477 GI:39883902
 VERSION CK297477
 KEYWORDS
 SOURCE
 ORGANISM

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT
 Other ESTs: EST760192
 Contact: Robin Buell
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850, USA
 Email: potato-array@igr.org
 Clones can be requested from TIGR via potato@tigr.org
 Seq primer: ATT TAG GTG ACA CTA TAG.

FEATURES

source

1. .853
 /organism="Nicotiana benthamiana"
 /mol_type="mRNA"
 /db_xref="taxon:4100"
 /clone="NBMD912"
 /tissue_type="biotic and biotic stress-treated leaves,
 callus tissue and root tissue"
 /lab_host="DH10B-Tona"
 /clone_lib="Nicotiana benthamiana mixed tissue cDNA
 library, normalized, full-length"
 /note="Vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI;
 supplier: RNA was isolated from Nicotiana benthamiana
 tissues that include callus, roots from liquid culture
 grown plants, heat-stressed leaves (38 C, 3 hr and 6 hr),
 cold-stressed leaves (5 C 3 hr, 6hr), and pathogen
 challenged leaves (Pseudomonas syringae pv tomato 12 hr;
 Xanthomonas campestris pv campestris 12 hr, 18hr;
 Pseudomonas syringae pv phaseolicola 18hr, and Xanthomonas
 campestris pv vesicatoria 18hr). RNA was isolated from
 these tissues and pooled in approximately equal molar
 amounts."

ORIGIN

Query Match 10.7%; Score 34.6; DB 14; Length 853;
 Best Local Similarity 47.5%; Pred. No. 25; Indels 0; Gaps 0;
 Matches 103; Conservative 0; Mismatches 114; Indels 0;

Fukuda, S., Hashizume, W., Hayashida, K., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kawai, J., Kojima, Y., Kondo, S., Kono, H., Koya, S., Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Saito, R., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M., Takeda, Y., Waki, K., Watanishi, A., Muramatsu, M. and Hayashizaki, Y. Direct Submision

Computational Analysis of Full-length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)

Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)

RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)

Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Tissues were provided by Dr. John Todd (Dept. of Medical Genetics Wellcome Trust Centre for Molecular Mechanisms in Disease Wellcome Trust/MRC building Addenbrookes Hospital Cambridge) whose assistance we gratefully acknowledge.

Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

FEATURES

source

1. 728
Location/Qualifiers
/organism="Mus musculus"
/mol_type="mRNA"
/strain="MOP"
/db_xref="taxon:10090"
/clone="F830034E14"
/issue_type="activated spleen"
/clone_lib="RIKEN full-length enriched, activated spleen"

ORIGIN

Query Match 10.7%; Score 34.8; DB 13; Length 728;
Best Local Similarity 55.9%; Pred. No. 20;

Matches 66; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

QY 188 CGATGATCTGAAGACTCTTACCTTGTGCTCTATTCGCTGCGATTCAGTGTGGCGC 247
DB 244 CGACGAGCTGATCTTTATCCCTTGTGTAAGCTGCTGTTAGACAGAGCTGGTGC 303
QY 248 ATTTCTTGGCGGTCCAGCAGCCTTGTGAGCGTTGGCGCTGGACACATCGTTTGT 305
DB 304 AACATCTGTGTCCAGCCCTCTCTTGTGCGTGTGTCTAAGCCATGCTTTGT 361

RESULT 9

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

B1144284 808 bp mRNA 1linear EST 05-JUL-2001
602907255F1 NCI_CGAP_Kid14 Mus musculus cDNA clone IMAGE:5064373
5', mRNA sequence.
B1144284
B1144284.1 GI:14604285
EST.
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgaps-remail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
CDNA Library Preparation: Life Technologies, Inc.

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
<http://image.lnl.gov>
Plate: L14M1173 row: m column: 14
High quality sequence start: 6
High quality sequence stop: 793.
Location/Qualifiers

FEATURES

source

1. 808
Location/Qualifiers
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:5064373"
/db_xref="taxon:10090"
/clone_lib="NCI_CGAP_Kid14"
/note="Organ: Kidney; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.75 kb. Constructed by Life Technologies. Note: this is a NCI_CGAP Library."

ORIGIN

Query Match 10.7%; Score 34.8; DB 12; Length 808;
Best Local Similarity 53.7%; Pred. No. 21;

Matches 72; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

QY 133 TGAATGGACAGAAATCCTTGCATTTTGACCGCATCAACGTTTGGACAGATGAGATA 192
DB 523 TGGAGGCTTCATGCTTCAACTTATGTGACGACCACTGGACATCAACAGACATG 582
QY 193 GATCTGAAGACTCTTACCTTGTGCTCATTCGCTTGCATTCAGTGTGGCGCATCTT 252
DB 583 GTTCAGAGATAGAAAGCTTGGGTTTCAAGCTTGTGTGCTCATGTAGATGCGCTGTA 642
QY 253 CTGGCGGTGACG 266
DB 643 CTGGCAATAGGCG 656

RESULT 10

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BX387771 1201 bp mRNA 1linear EST 08-MAY-2003
BX387771 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA clone CS0D1052YB18 5-PRIME, mRNA sequence.
BX387771
BX387771.1 GI:30460721
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 4682.r For more information about this cluster, see
<http://www.genoscope.cns.fr/>
cgi-bin/cluster.cgi?seq=CS0A1052DA09QPlcluster=4682.r. Contact :
Feng Huang Email : Fluang@litech.com URL :
<http://fulllength.invitrogen.com/> Invitrogen Corporation 1600
Paradise Avenue Genoscope sequence ID : CS0A1052DA09QPl.
Location/Qualifiers
1. 1201
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"

Tue Apr 20 06:47:20 2004

us-10-608-504-4.rst

Page 4

ORIGIN

Query Match 10.9%; Score 35.2; DB 28; Length 910;
Best Local Similarity 57.1%; Pred. No. 17;
Matches 64; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBTAS30007"
/clone_1ib="ZM 0.6 1.0 KB"
/note="Vector: pCR4-TOPO, Site 1: EcoRI; 0.6-1.0 kb high
cot selected genomic DNA library"

QY 192 AGATCTGAAGACTTAACTTTGCTGCTACCTGCGATTAACAGTGGCGGCACT 251
DB 272 AATCTGACGGCTCAATCTTTTCCCAATCCGACATCGGACGAGGCGCGCC 331
QY 252 TCTTGCGCGTCCGACGACCTTGTGAGCGTTGCGCTGCGACCATGTTTT 303
DB 332 TGGTCCGCGGATCTCTGCTGCGACGACCGCATGCTGCGCGCGTTT 383

RESULT 7 1101 bp DNA linear GSS 04-JUN-1999
CNS00BD8
LOCUS Drosophila melanogaster genome survey sequence TET3 end of BAC #
DEFINITION BAC23H11 of RPCI-98 library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
ACCESSION AL056828.1 GI:4937496
VERSION GSS.
KEYWORDS Drosophila melanogaster (fruit fly)
SOURCE Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachyera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 1101)
REFERENCE Genoscope.
AUTHORS Direct Submission
TITLE Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
JOURNAL BP 191 91006 EVRY cedex - FRANCE (E-mail : sequef@genoscope.cns.fr
- Web : www.genoscope.cns.fr)

COMMENT
The BAC end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using the BACs. For further information
please see <http://www.fruitfly.org> The BDGP Drosophila
melanogaster BAC library was prepared by Kazuo Oosawa and
Aaron Mammeter in Peter de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPCI-98 and was constructed by partial
EcoRI digestion of Drosophila DNA provided by the BDGP from the
isogenic strain y2; cn bw sp, the same strain used for the BDGP's
P1 and EST libraries. A more detailed description of the library
and how to order individual BAC clones, the entire library, or
filters for hybridization from the BACPAC Resource Center can be
found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES

Source
1. 1101
Location/Qualifiers
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACR23H11"
/clone_1ib="RPCI-98"
/note="end : TET3"

ORIGIN

Query Match 10.9%; Score 35.2; DB 29; Length 1101;
Best Local Similarity 24.4%; Pred. No. 18;
Matches 62; Conservative 81; Mismatches 111; Indels 0; Gaps 0;

QY 29 TTGTTGCGACATGTGACAGTATTAATCTTTGGCGCTCCGCGGCTTCCTTAATCC 88
DB 847 TTGGGKTGGTGGGTAADTGTATTAATTTTBTBKTBTBTBNAATKWTASSTTTT 906

QY 89 TTAGCCCTTACGTATCACTATTTTGCGCAAAATGGCCATGTGATCCAGCAGGA 148
DB 907 WSGAGGCTGTBKBSKGTSTSVGTTTWSGSAVAASSTTTTWTSTBTKABBA 966
QY 149 TCTTGCATTTTGAACCATCAACGTTTGGCAGCAATGCAATGATGCAAGACTTAA 208
DB 967 ATSAHTTASDPTWAABSSKTSVBSBSSDMGRTASGATTTTCTTTTWTWW 1026
QY 209 CTTTGTCTCATTTGCGCTTGGCATTAACGTGTGCGCATCTTCTGCGGTCAGCA 268
DB 1027 TATTTTNSBMBWBAATTTTDTAAATWAATBTSNNSNMNNTTBTSSRAATRAABA 1086
QY 269 CTTTGTGAGCGTT 282
DB 1087 SGTBGTBBSVATB 1100

RESULT 8 728 bp mRNA linear EST 17-DEC-2002
BY765711
LOCUS Drosophila melanogaster genome survey sequence TET3 end of BAC #
DEFINITION BAC23H11 of RPCI-98 library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
ACCESSION BY765711.1 GI:27203910
VERSION EST.
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 728)
REFERENCE Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S.,
Nikaido, I., Oosato, N., Saito, R., Suzuki, H., Yamataka, I.,
Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A.,
Schonbach, C., Gotohori, T., Baldarelli, R., Hill, D.P., Balt, C.,
Hume, D.A., Quackenbush, J., Schriml, L.M., Kanapi, A., Matsuda, H.,
Baturov, S., Beisel, K.M., Blake, J.A., Brad, D., Brucic, V.,
Ciothia, C., Corbett, L.E., Cousins, S., Dalla, E., Dragani, T.A.,
Fletcher, C.F., Forrest, A., Frazer, K.S., Gaasterland, T.,
Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S.,
Guernich, S., Hirokawa, N., Jackson, I.J., Jarvis, E.D., Kanai, A.,
Kawai, H., Kawasawa, Y., Kedzierski, R.M., King, B.L., Konagaya, A.,
Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglott, D.R.,
Matsuda, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T.,
Numata, K., Okido, I., Pavan, W.J., Pettes, G., Pesole, G.,
Petrovsky, N., Pillai, R., Pontius, J.V., Qi, D., Ramachandran, S.,
Ravasi, T., Reed, J.C., Reed, D.J., Reid, J., Ring, B.Z., Ringwald, M.,
Sandelin, A., Schneider, C., Semple, C.A., Setou, M., Shimada, K.,
Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale, R.D., Tomita, M.,
Verardo, R., Wagner, L., Wahlstedt, C., Wang, Y., Watanabe, Y.,
Wells, C., Wilmberg, L.G., Wymshaw-Boris, A., Yanagisawa, M., Yang, I.,
Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Cantinci, P.,
Hayatsu, N., Hirozane-Kishikawa, T., Komori, H., Nakamura, M.,
Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K.,
Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imutani, K., Ishii, Y.,
Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K.,
Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E.S.,
Rogers, J., Birney, E. and Hayashizaki, Y.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)

TITLE
JOURNAL
MEDLINE
PUBMED
COMMENT
Contact: Yoshinori Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@sc.riken.go.jp/
URL: <http://genome.gsc.riken.go.jp/>
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P.,

LOCUS CF338060 475 bp mRNA linear EST 18-AUG-2003
 DEFINITION JMT--08-M05.b1 AtUMT-overexpressing transgenic rice plasmid cDNA
 library (JMT) Oryza sativa cDNA clone JMT--08-M05, mRNA sequence.
 ACCESSION CF338060
 VERSION CF338060.1 GI:33824491
 KEYWORDS EST.
 SOURCE Oryza sativa
 ORGANISM Oryza sativa
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Ehrhartoideae; Oryzaceae; Oryza.
 REFERENCE 1 (bases 1 to 475)
 Kim,J.S., Jun,K.M., Cheong,P.J., Kim,M.J., Lee,T.H., Shin,Y.C.,
 Song,S.I., Kim,J.K., Kim,Y.-K. and Nahm,B.H.
 Large-scale Sequencing Analysis of Rice ESTs
 Unpublished (2003)
 TITLE JOURNAL
 COMMENT Contact: Nahm B.H.
 Genomics and Genetics Institute, Greengene Biotech Inc.; Division
 of Bioscience and Bioinformatics, Myongji University
 Yongin, Kyeonggi, Korea
 Tel: 82 31 330 6193
 Fax: 82 31 321 6355
 Email: bhnahm@bio.com, bhnahm@bio.myongji.ac.kr.

FEATURES
 source
 1..475
 /organism="Oryza sativa"
 /mol_type="mRNA"
 /cultiVar="Nackdong"
 /db_xref="taxon:4530"
 /clone="JMT--08-M05"
 /tissue_type="leaf"
 /dev_stage="14 days after germination"
 /lab_host="E.coli DH10B"
 /clone_lib="AtUMT-overexpressing transgenic rice plasmid
 cDNA library (JMT)"
 /note="Vector: pCR4-TOPO, Site 1: EcoRI; Oligo-capped mRNA
 was reverse transcribed and then used for PCR. mRNA was
 prepared from Arabidopsis vasominate Carboxyl
 methyltransferase overexpression line."

ORIGIN
 Query Match 10.9%; Score 35.2; DB 14; Length 475;
 Best Local Similarity 50.6%; Pred. No. 13; Indels 0; Gaps 0;
 Matches 85; Conservative 0; Mismatches 83;

Qy 51 CATTACTTTGGCGTCCGCGGCTTCGTTCTTAATCCTTAAGCCCTACGTGAATCA 110
 Db 255 CTTCTTTCTGGGCTCCCGGGCTTCTGCTAGCCTTCTTAACGCCGCTTGGCGGCG 196
 Qy 111 ATTGTGGGCAAAATGGGATCTGTGATCCAGCGAGATCCTTGCCATTGGACCGCATC 170
 Db 195 CGGCGCTTGGCGGCTCTACTAGCTTGCGACGCCGCTTGGGAACTCCGCGAGTTT 136
 Qy 171 AACGTTTGGCAGCATGCGATGATCTGAAGACTCTTAACCTTTGGTCT 218
 Db 135 CACGCGCTTCAAGAGAAGCATCTTCTGATGCTTCTGCTGCTGTTCT 88

RESULT 5
 B2720836 786 bp DNA linear GSS 24-FEB-2003
 LOCUS PUCFW05TD.ZM.0.6.1.0 KB Zea mays genomic clone ZM8MTa137A09,
 DEFINITION genomic survey sequence.
 ACCESSION B2720836
 VERSION B2720836.1 GI:28512409
 KEYWORDS GSS.
 SOURCE Zea mays
 ORGANISM Zea mays
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
 clade; Panicoidae; Andropogonaeae; Zea.
 REFERENCE 1 (bases 1 to 786)
 Whitelaw,C.A., Quackenbush,J., Van Aken,S., Uterback,T.,

Renick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
 Bennettzen,J.
 Maize Genomics Consortium
 Unpublished (2003)
 TITLE JOURNAL
 COMMENT Contact: Cathy Whitelaw
 TIGR
 9712 Medical Center Drive, Rockville, MD 20850, USA
 Tel: 301-838-5843
 Fax: 301-838-0208
 Email: whitelaw@tigr.org
 Seq primer: TF
 Class: sheared ends.

FEATURES
 source
 1..786
 /organism="Zea mays"
 /mol_type="genomic DNA"
 /strain="B73"
 /db_xref="taxon:4577"
 /clone="ZM8MTa137A09"
 /clone_lib="ZM.0.6.1.0 KB"
 /note="Vector: pCR4-TOPO, Site 1: EcoRI; 0.6-1.0 kb high
 COT selected genomic DNA library"

ORIGIN
 Query Match 10.9%; Score 35.2; DB 28; Length 786;
 Best Local Similarity 53.7%; Pred. No. 16; Indels 0; Gaps 0;
 Matches 73; Conservative 0; Mismatches 63;

Qy 189 GATAGATCGAAGACTCTAACCTTGTCTCAATGCGCGTGGAGTACAGTGGTGGCGCA 248
 Db 480 GAAGCAATCAAAATTTGTGTCTAGTCTCAAGGTGGTGGCAATGTGTGGCTTA 539
 Qy 249 TCTTCTGGCGGTGACGACCTTGTGAGCGTTGGCGCTGACACCATCTTTTGTGG 308
 Db 540 TCATTTAGGTGTCATCTTGAATGTTATCATTTGGCCCCCAGCAATCTTGTGATCG 599
 Qy 309 ACTGTGATCTTTTC 324
 Db 600 TATGTTGCATGTGTC 615

RESULT 6
 CC362201 910 bp DNA linear GSS 16-MAY-2003
 LOCUS PHTK88TD.ZM.0.6.1.0 KB Zea mays genomic clone ZM8MTa1330007,
 DEFINITION genomic survey sequence.
 ACCESSION CC362201
 VERSION CC362201.1 GI:30831601
 KEYWORDS GSS.
 SOURCE Zea mays
 ORGANISM Zea mays
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
 clade; Panicoidae; Andropogonaeae; Zea.
 REFERENCE 1 (bases 1 to 910)
 Whitelaw,C.A., Quackenbush,J., Van Aken,S., Uterback,T.,
 Renick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
 Bennettzen,J.
 Maize Genomics Consortium
 Unpublished (2003)
 TITLE JOURNAL
 COMMENT Other GSSs: PHTK88TB
 Contact: Cathy Whitelaw
 TIGR
 9712 Medical Center Drive, Rockville, MD 20850, USA
 Tel: 301-838-5843
 Fax: 301-838-0208
 Email: whitelaw@tigr.org
 Seq primer: TF
 Class: sheared ends.

FEATURES
 source
 1..910
 /organism="Zea mays"
 /mol_type="genomic DNA"

Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCL-98 and was constructed by partial EcoRI digestion of *Drosophila* DNA provided by the BDSP from the isogenic strain Y2; cn bw sp, the same strain used for the BDSP-Pl and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or funds for hybridization from the BAC/CDSP Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES

```

1..935
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACR12G04"
/clone_1lb="RP11-98"
/note="end : Ter3"

```

ORIGIN

Query Match	11.7%;	Score 37.8;	DB 29;	Length 995;
Best Local Similarity	34.0%;	Pred. No. 2,9;		
Matches	54;	Conservative	41;	Mismatches 64;
				Indels 0;
				Gaps 0

Cy	16	CCGATCAAGCGTTTCGAGAAATGCATGTACATAAGCTTAACTTCATTGGTCCATTG	2233
Dd	872	CAGTMMTTAASGAGCGGTGCTKTBMAAAGAASASASTTGVTSTSGCTTTMTTYTAAT	8133
Ox	224	CCGTTGGGATTACAAGTAGGCGGCACATCTCTTGAGCGGTGAACGACACTGTGTGACCTTG	2833
Dd	812	CSKTTTSBPTMGKACSCHMTTTSASMSAWTCGTGCKKKGGAGSRRAPAAAGGCGCAKSKYTB	7533
Ox	284	GCGCTGGCACCATTCGTTTTTGTGTGACACTGGTGAATCTTT	322
		::::	
Dd	752	SSGRTGATKRYSTSAKKTGGTRGGSTGADPAMTCTTT	714

RESULT 2			
BZ557685			
BZ557685			
LOCUS			
DEFINITION			
	BZ557685	1353 bp	DNA
	pac1-60_663..81		linear GSs 17-DEC-2002
	pac1-60_663, genomic		<i>Pseudomonas aeruginosa</i> genomic clone
	pac1-60_663, genomic		survey sequence.

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

Source

```

/organism="Pseudomonas aeruginosa"
/mol_type="genomic DNA"
/strain="1-60"
/db_xref="taxon:287"
/cclone="pacal-60-663"
/cclone_1b="pacal-60"
/note="clinical isolate 1-60 whole genomic shotgun library."

```

ORIGIN

Query Match	11.0%;	Score 35.8;	DB 28;	Length 1353;
Best Local Similarity	57.7%;	Pred. No. 13;		
Matches 64;	Conservative 0;	Mismatches 47;	Indels 0;	Gaps 0;

QY 204 TCTAACTTTGGCTCATATGGCGGTGGGCATTAACAATGTTGGCAGATCTCTGAGCGGCTG 28

Db 86 TATAGCCCTTGCTCTCGTGGGAGTGAAGTCGCCCTGGCGGCGCTTGAATCTTGGCTGATCC 145

QY 264 AACCACCTTGTGAGCGTGGGCGTGGACACCATCTTTTGTGACATGGT 314

Db 146 GCCGCAAGGAGCTTCTGATCGGAGGCTCTGTCGGCGCTTTCGGTGGCCAGGT 196

RESULT 3	CF302597/c	471 bp	mRNA	linear	EST 15-AUG-2003
LOCUS	7L5AF--08-E15.b1	Rice	leaf	plasmid	cd DNA library II (7LEAF) Oryza
DEFINITION	bativa cdna clone 7LEAF--08-E15,		mRNA	sequence:	

SOURCE ORGANISM
Oryza sativa
Oryza sativa
Embryos: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehretaceae; Oryzae; Oryza.

COMMENT

FEATURES

Genomics and Bioinformatics, Myongji University
of Bioscience and Bioinformatics, Myongji University
Yongin, Kyoeonggi, Korea
Tel: 82 31 330 6193
Fax: 82 31 321 6355
Email: bhnam@gbic.com, bhnam@bio.myongji.ac.kr.
Location/Qualifiers

FEATURES

```

/organism="Oryza sativa"
/mol_type="mRNA"
/cultivar="Nackdong"
/db_xref="taxon:4530"
/clone="7LEAF--08-E15"
/tissue_type="leaf"
/dev_stage="7 days after germination"
/lab_host="E.coli DH10B"
/clone_lib="Rice leaf plasmid cDNA library II (7LEAF)"
/note="Vector: pCR4-TOPO. Site 1: EcoRI; mRNA was capped
with oligoribonucleotides and then used as templates for
RT-PCR."

```

	Query Match	Best Local Similarity	Score 35.2;	DB 14;	Length 471;
	Matches	85; Conservative	50.6%;	Pred. No. 13;	Indels 0; Gaps 0
Qy	51	CATTACTTTGGCGCTCCGGCGGATTCCTTTCTTAATCCTTAAGCCCGCCAGTGAATACAA	110		
Db	224	CCCTCTCTTGGGCGCTCCGGGCGTTCTGCTACGCCCTTTTGAAGCGCGCCCTTGGCGGCG	165		
Qy	111	ATTGTGGCAAAATGGCGATGTGATGCGACAGGAATCCTTGCATTTGAACCGCAATC	170		
Db	164	CGGCGTTTGGCGCGCTCTCACTAGCTTGGCAGCGCGCTCCTTGGCAATCGCGCAGATT	105		
Qy	171	AAGCTTCCAGCAATGCGATAGTCTGAAGACTCAACCTTTGACT	218		
Db	104	CAGGCGCTTCAAGGAAGCATTCTCTGATGCTCTCTGCTCTGTTCT	57		

RESULT 4
CF338060/c

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 17, 2004, 06:48:18 ; Search time 1049.18 Seconds
(without alignments)
9221.792 Million cell updates/sec

Title: US-10-608-504-4

Perfect score: 324
Sequence: 1 atgacaactgattctctctcctcgttgcgttccttc 324

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: em_estba:*
2: em_estba:*
3: em_estba:*
4: em_estba:*
5: em_estba:*
6: em_estba:*
7: em_estba:*
8: em_estba:*
9: gb_est1:*
10: gb_est2:*
11: gb_est3:*
12: gb_est4:*
13: gb_est5:*
14: gb_est6:*
15: em_estba:*
16: em_estba:*
17: em_gss_hum:*
18: em_gss_hum:*
19: em_gss_hum:*
20: em_gss_hum:*
21: em_gss_hum:*
22: em_gss_hum:*
23: em_gss_hum:*
24: em_gss_hum:*
25: em_gss_hum:*
26: em_gss_hum:*
27: em_gss_hum:*
28: gb_est1:*
29: gb_est2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	37.8	11.7	995	29	CNS005NG
C 2	35.8	11.0	1353	28	B2557685
C 3	35.2	10.9	471	14	CF302597
C 4	35.2	10.9	475	14	CF338060

Result No.	Score	Query Match	Length	ID	Description
5	35.2	10.9	786	28	B2720836
6	35.2	10.9	910	28	CC362201
7	35.2	10.9	1101	29	CNS005NG
8	34.8	10.7	728	13	BY765711
9	34.8	10.7	808	12	B1144284
10	34.8	10.7	1201	13	BK387771
11	34.6	10.7	413	14	CB239541
12	34.6	10.7	853	14	CK297477
13	34.6	10.7	890	14	CK288648
14	34.6	10.7	922	14	CK295854
15	34.6	10.7	936	14	CK295424
16	34.6	10.7	951	14	CK292130
17	34.4	10.6	447	29	CNS02NHL
18	34.4	10.6	1201	13	BK356376
19	34.2	10.6	1062	28	B2557686
20	34.2	10.5	423	28	B44402
21	33.8	10.4	1088	29	CNS0468Q
22	33.6	10.4	833	13	BUS98308
23	33.4	10.3	580	12	BM309296
24	33.2	10.2	478	14	H65669
25	33.2	10.2	586	14	CD091362
26	33.2	10.2	859	10	BF341287
27	33.2	10.2	1044	28	BZ533873
28	33.2	10.2	569	12	BM100089
29	33.2	10.2	720	29	CC721864
30	33.2	10.2	761	29	CG084816
31	33.2	10.2	796	29	CC684833
32	33.2	10.2	818	28	CC005623
33	33.2	10.2	958	29	CC684842
34	32.8	10.1	390	10	AM456947
35	32.8	10.1	399	10	AM491785
36	32.8	10.1	416	10	AM457742
37	32.8	10.1	463	10	AM489067
38	32.8	10.1	469	10	BF460793
39	32.8	10.1	466	12	B1135202
40	32.8	10.1	496	12	B1134847
41	32.8	10.1	536	12	B1901653
42	32.8	10.1	565	12	B1901912
43	32.8	10.1	602	14	BF207637
44	32.8	10.1	638	14	CF582833
45	32.8	10.1	3361	11	AK083916

ALIGNMENTS

RESULT 1
CNS005NG/LOCUS
DEFINITION
Drosophila melanogaster genome survey sequence (EST3 end of BAC # BACR12G04 of RPLT-98 library from Drosophila melanogaster (fruit fly)), genomic survey sequence.
ACCESSION
AL060428
VERSION
AL060428.1 GI:4943359
KEYWORDS
GSS.
SOURCE
Drosophila melanogaster (fruit fly)
ORGANISM
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE
1 (bases 1 to 995)
Genoscope.
Direct Submission
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see <http://www.fruitfly.org> The BDGP Drosophila melanogaster BAC library was prepared by Kazuo Ooe and Aaron Mammose in Pletier de Jong's laboratory in the Department of

COMMENT

QY 178 CCAGCATGATAGATCTGAAGACTTAACCTTTGTCATTCGCGTTGCGATTACA 237
DB 15721 GGCTCCACGCTGGGGAAGAGGCTCTGAAGCAGACGACCATCTGCTGCGGTGTA 15780
QY 238 GTGGTGGCGCATCTTTCTTGCGC 259
DB 15781 GTGGTACTCTATCTTCTGATG 15802

RESULT 14

US-10-796-280-12378
Sequence 12378, Application US/10796280
GENERAL INFORMATION:
APPLICANT: CARGILL, Michele et al.
TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
FILE REFERENCE: CL001510
CURRENT APPLICATION NUMBER: US/10/796,280
CURRENT FILING DATE: 2004-03-10
NUMBER OF SEQ ID NOS: 68533
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 12378
LENGTH: 234309
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(234309)
OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1-
US-10-796-280-12378

Query Match 8.8%; Score 28.4; DB 6; Length 234309;
Best Local Similarity 52.5%; Pred. No. 81;
Matches 62; Conservative 0; Mismatches 56; Indels 0; Gaps 0;

QY 206 TAACCTTTGTCATTCGCGTTGCGATTACAGTGGCGCATCTTTGCGGTCGAC 265
DB 62038 TCACCTCAGGCGCTTTAGCCTGAAGAGCGTGCGCCCTACATCTGCTTTAGTTTCA 62097
QY 266 GCACTTTGTTGAGCGTTGGCGCTGGCAGCATGCTTTTGTGACTGGTAATCTTTT 323
DB 62098 GTCCCTGTGGTAAGTTGGTAAGGCCCTTTCGCTGTGTGTAAAGAGGGGTTT 62155

RESULT 15

US-10-767-701-27384/C
Sequence 27384, Application US/10767701
GENERAL INFORMATION:
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(53535) B
CURRENT APPLICATION NUMBER: US/10/767,701
CURRENT FILING DATE: 2004-01-29
NUMBER OF SEQ ID NOS: 63128
SEQ ID NO 27384
LENGTH: 490
TYPE: DNA
ORGANISM: Sorghum bicolor
FEATURE:
OTHER INFORMATION: Clone ID: 6858297
US-10-767-701-27384

Query Match 8.7%; Score 28.2; DB 6; Length 490;
Best Local Similarity 51.2%; Pred. No. 5.8;
Matches 66; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

QY 192 AGATCTGAAGACTTAACCTTTGTCATTCGCGTTGCGATTACAGTGGCGCATCT 251
DB 282 AGGTGTACAGTGTAGCAGTTGCTAGCTGCCAGAGCGAGCGAGTGAGCCAGCGCC 223

QY 252 TCTTGGCGGTGACAGCAGCACTTTGTTGAGCGTTGGCGCTGGCACCACCATGTTTTTGGACT 311
DB 222 AGTTGGCGGTGAGAGGTGACCGACTGGGTGGCGGTGTGCAACCAAGACTTGTGTGACT 163
QY 312 GTGATCT 320
DB 162 GGGCAACT 154

Search completed: April 17, 2004, 20:13:00
Job time : 39.4407 secs

QY 143 CAGGAATCTTCCATTTTGACCCGATCAGCTTTCGAGCAATGCAATGATCTGAGA 202
 DB 284 CCGGATTCCTGGCGGAGCTATACCGACGCTATTAACCGCATAGCTTGCCCTGCCACGG 225
 QY 203 CTCTAATCTTGGTCTCATTTGCGGTTGCGATTACAGTGGTGGCGCATCTTTCGCGGTC 262
 DB 224 CACTTAACCGGTACGGGTATTTGCGCTTCCGAGCAACTGATTTGCGCTTCAAGTGGCGATG 165
 QY 263 GAGCGACCTTGTG 276
 DB 164 TTCTGATCATGATG 151

RESULT 10
 US-10-108-605A-248/C
 ; Sequence 248, Application US/10108605A
 ; GENERAL INFORMATION:
 ; APPLICANT: Broadus, Julie
 ; APPLICANT: Steam, Lynn
 ; APPLICANT: Bachmann, Jane
 ; APPLICANT: Kander, Kim
 ; TITLE OF INVENTION: NUCLEIC ACID SEQUENCES FROM DROSOPHILA MELANOGASTER THAT ENCODE
 ; FILE REFERENCE: 3113B
 ; CURRENT APPLICATION NUMBER: US/10/108,605A
 ; CURRENT FILING DATE: 2002-03-27
 ; PRIOR APPLICATION NUMBER: US 09/761,142
 ; PRIOR FILING DATE: 2001-01-16
 ; PRIOR APPLICATION NUMBER: US 60/176,418
 ; PRIOR FILING DATE: 2000-01-14
 ; NUMBER OF SEQ ID NOS: 361
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 248
 ; LENGTH: 8577
 ; TYPE: DNA
 ; ORGANISM: Drosophila melanogaster
 US-10-108-605A-248

Query Match 8.8%; Score 28.4; DB 6; Length 8577;
 Best Local Similarity 60.3%; Pred. No. 18;
 Matches 47; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

QY 209 CCTTGATCTTCATTCGCGTTGCGATTACAGTGGTGGCGCATCTTTCGCGGTCCAGCA 268
 DB 83 CCGTGTGTTGTTGGCGCCACTGCTACTGTGTCTTCTTGGCAGGCGACAC 24
 QY 269 CCTTGTGAGCGTTGGCG 286
 DB 23 CCGTTTCAACGGCGGTG 6

RESULT 11
 US-10-796-280-12417/C
 ; Sequence 12417, Application US/10796280
 ; GENERAL INFORMATION:
 ; APPLICANT: CARGILL, Michele et al.
 ; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
 ; FILE REFERENCE: C1001510
 ; CURRENT APPLICATION NUMBER: US/10/796,280
 ; CURRENT FILING DATE: 2004-03-10
 ; NUMBER OF SEQ ID NOS: 68533
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 12417
 ; LENGTH: 20618
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-10-796-280-12417

Query Match 8.8%; Score 28.4; DB 6; Length 20618;
 Best Local Similarity 51.6%; Pred. No. 27;
 Matches 65; Conservative 0; Mismatches 61; Indels 0; Gaps 0;

QY 65 TCCGGGGGGTTCCGTTCTTAATCCTTAAGCCCTAGCATCATCAATTTGGGCAAAA 124
 DB 8143 TCCCATATTTCCACTGTTTATCTTTGGCTCAAGCAGACATCACTCTCTTGATTTA 8084
 QY 125 TGGCGATGTGATGACGAGCAAGAAATCTTGCAATTTTGACCGCATCAACGTTTCGAGCA 184
 DB 8083 TGCTGCTCTACTGCTTCATCTCCTGCTGCTGCTGCTTCCCACTATTTCTCACA 8024
 QY 185 ATGGCA 190
 DB 8023 AGGCA 8018

RESULT 12
 US-10-796-280-12418
 ; Sequence 12418, Application US/10796280
 ; GENERAL INFORMATION:
 ; APPLICANT: CARGILL, Michele et al.
 ; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
 ; FILE REFERENCE: C1001510
 ; CURRENT APPLICATION NUMBER: US/10/796,280
 ; CURRENT FILING DATE: 2004-03-10
 ; NUMBER OF SEQ ID NOS: 68533
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 12418
 ; LENGTH: 46075
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-10-796-280-12418

Query Match 8.8%; Score 28.4; DB 6; Length 46075;
 Best Local Similarity 51.6%; Pred. No. 39;
 Matches 65; Conservative 0; Mismatches 61; Indels 0; Gaps 0;

QY 65 TCCGGGGGGTTCCGTTCTTAATCCTTAAGCCCTAGCATCAATTTGGGCAAAA 124
 DB 3888 TCCCATATTTCCACTGTTTATCTTTGGCTCAAGCAGACATCACTCTCTTGATTTA 3947
 QY 125 TGGCGATGTGATGACGAGCAAGAAATCTTGCAATTTTGACCGCATCAACGTTTCGAGCA 184
 DB 3948 TGCTGCTCTACTGCTTCATCTCCTGCTGCTGCTGCTTCCCACTATTTCTCACA 4007
 QY 185 ATGGCA 190
 DB 4008 AGGCA 4013

RESULT 13
 US-10-796-280-12452
 ; Sequence 12452, Application US/10796280
 ; GENERAL INFORMATION:
 ; APPLICANT: CARGILL, Michele et al.
 ; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
 ; FILE REFERENCE: C1001510
 ; CURRENT APPLICATION NUMBER: US/10/796,280
 ; CURRENT FILING DATE: 2004-03-10
 ; NUMBER OF SEQ ID NOS: 68533
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 12452
 ; LENGTH: 69359
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-10-796-280-12452

Query Match 8.8%; Score 28.4; DB 6; Length 69359;
 Best Local Similarity 50.0%; Pred. No. 47;
 Matches 71; Conservative 0; Mismatches 71; Indels 0; Gaps 0;
 QY 118 GGCATAATGCGGATGTGATGCCAGAGAAATCTTGCAATTTTGACCGCATCAACGTTT 177
 DB 15661 GGCATAATGCGGATGTGATGCCAGAGAAATCTTGCAATTTTGACCGCATCAACGTTT 15720

QY 108 ACAATTGTGGCAAAATGCGATGTGATGCCACGAGA 147
DB 48937 ACAAGTTGAATTCNAATGCCAAGGTGATGCGCTTGGGA 48898

RESULT 6
US-10-650-609-1/c

Sequence 1, Application US/10650609
GENERAL INFORMATION:
APPLICANT: Costa, Robert
APPLICANT: Wang, Xinde
APPLICANT: Tan, Yonjun
APPLICANT: Kalinichenko, Vladimir
APPLICANT: Krupczak-Holts, Katherine
APPLICANT: Wang, I-Ching
APPLICANT: Major, Michael
TITLE OF INVENTION: METHODS OF TREATING AGE-RELATED DEFECTS AND DISEASES
FILE REFERENCE: 02-1040-A
CURRENT APPLICATION NUMBER: US/10/650,609
CURRENT FILING DATE: 2003-08-28
PRIOR APPLICATION NUMBER: 60/406582
PRIOR FILING DATE: 2003-08-28
PRIOR APPLICATION NUMBER: 60/426068
PRIOR FILING DATE: 2002-11-13
NUMBER OF SEQ ID NOS: 3
SOFTWARE: PatentIn version 3.0
SEQ ID NO 1
LENGTH: 2737
TYPE: DNA
ORGANISM: Homo sapiens
US-10-650-609-1

Query Match
Best Local Similarity 8.8%; Score 28.6; DB 6; Length 2737;
Matches 55; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

QY 61 GCGTCGCGGCGGCTTCCGTTCTTAATCCTTAAGCCCTACGTAATCAATTTGTGGGC 120
DB 1109 GAGCGCTGTGATCGGTCGCTTCTGCTGCTTAACACCTGTCATGCAAGTACGCGTT 1050
QY 121 AAATGCGCATGTGATGCCACAGAAATCCTTGCCATT 159
DB 1049 GGCACCTGGGATGATGCTCCAGAGAGACCTTGCCATT 1011

RESULT 7
US-10-451-467A-545/c

Sequence 545, Application US/10451467A
GENERAL INFORMATION:
APPLICANT: CONTRERAS, ROLAND HENRI
APPLICANT: EBERHARDT, INES
APPLICANT: LUYTEN, WALTER HERMAN MARIA LOUIS
APPLICANT: REEKMAN, RIEKA JOSEPHINA
TITLE OF INVENTION: BAX-RESPONSIVE GENES FOR DRUG TARGET IDENTIFICATION IN
FILE REFERENCE: JAB-1667
CURRENT APPLICATION NUMBER: US/10/451,467A
CURRENT FILING DATE: 2003-06-19
PRIOR APPLICATION NUMBER: EP 00870318.3
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: EP 01870002.1
PRIOR FILING DATE: 2001-01-04
PRIOR APPLICATION NUMBER: EP 01870003.9
PRIOR FILING DATE: 2001-01-09
NUMBER OF SEQ ID NOS: 732
SOFTWARE: PatentIn version 3.1
SEQ ID NO 545
LENGTH: 3042
TYPE: DNA
ORGANISM: Candida albicans
US-10-451-467A-545

Query Match
Best Local Similarity 8.8%; Score 28.6; DB 6; Length 3042;
Matches 49; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

QY 9 TGAATTCCTCGTATTCCTGTTGTGTCGAGATGTGCAATTAATTTGCGCTCG 68
DB 956 TGTGCTCTTTCATCTCTGTTGTTCAATTTGTTCTTTCTTTGATGATTT 897
QY 69 GCGGTCGCTTCTTAATCCTTA 91
DB 896 TGTGTTCCAAATTTGAATTTTA 874

RESULT 8
US-10-767-701-21673/c

Sequence 21673, Application US/10767701
GENERAL INFORMATION:
APPLICANT: KOVALIC, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(5353)B
CURRENT APPLICATION NUMBER: US/10/767,701
CURRENT FILING DATE: 2004-01-29
NUMBER OF SEQ ID NOS: 63128
SEQ ID NO 21673
LENGTH: 312
TYPE: DNA
ORGANISM: Sorghum bicolor
FEATURE:
OTHER INFORMATION: Clone ID: 13317221
US-10-767-701-21673

Query Match
Best Local Similarity 8.8%; Score 28.4; DB 6; Length 312;
Matches 59; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

QY 204 TCTAACCTTGTGCTCATTCGCTTGCGATTAACAGTGTGGCGCATCTTGTGGCGGTG 263
DB 261 TCTTATCTCTTCGCGCTTGCCTTCGCTGCTGCGCTTCTTCTCGGGAGAG 202
QY 264 ACGACCTTGTGAGCGTGGCGCTGACCAATCGTTTGTGACTG 313
DB 201 GCGCGAGGAGCGCGCGTGGCGCTGCGCGCGCGCATGGAGAGCGGG 152

RESULT 9
PCT-US04-02000-767/c

Sequence 767, Application PC/TUS0402000
GENERAL INFORMATION:
APPLICANT: Science Applications International Corporation
TITLE OF INVENTION: Method and System for Identifying Biological Entities in
FILE REFERENCE: 36609-183264 (SAIC0087-PCT)
CURRENT APPLICATION NUMBER: PCT/US04/02000
CURRENT FILING DATE: 2004-01-23
PRIOR APPLICATION NUMBER: US 60/441,745
PRIOR FILING DATE: 2003-01-23
PRIOR APPLICATION NUMBER: US 60/441,806
PRIOR FILING DATE: 2003-01-23
NUMBER OF SEQ ID NOS: 3242
SOFTWARE: PatentIn version 3.2
SEQ ID NO 767
LENGTH: 500
TYPE: DNA
ORGANISM: Escherichia coli
PCT-US04-02000-767

Query Match
Best Local Similarity 8.8%; Score 28.4; DB 1; Length 500;
Matches 68; Conservative 0; Mismatches 66; Indels 0; Gaps 0;

US-10-789-400-36
Commerce 36 A

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

Query Match	9.4%;	Score 30.4;	DB 6;	Length 13350;
Best Local Similarity	49.4%;	Pred. No. 4.5;		
Matches	79;	Mismatches	81;	Indels 0;
				Gaps 0;

14 TCTCCTGATTCCTTGTGTGCGAGTATGTGCACTTACTTTTGGCCCTCCGGGCGG 73

Db 1687 TCTCAGATAATGAGAGAGAGATGCGATCTTCAATCTTAACCTTGAAGAAGAGATA 1746

74 TTCGCTCTTAATCCCTTAAGCCCTACGTGAATCACAATTGTGTGGGCAAAATGGCGATGT 133

Db 1747 CTTGATCATTAAGCATTGAGGCCAGATTGGAATCATAGAGGAGAAATTAAGCATGATAT 1806

134 GGATGCCAGCAGGAACTCTGCCATTTTGACCGCATCAAC 173

1807 TAGGCTATTAGACACTCAACATTGCTACAGCAGGACC 1848

```

RESULT 3
US-10-767-471-10749/c
; Sequence 10749, Application US/10767471
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: RHEUMATOID ARTHRITIS, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO011505
; CURRENT APPLICATION NUMBER: US/10/767,471
; CURRENT FILING DATE: 2004-01-30
; NUMBER OF SEQ ID NOS: 50231
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10749
; LENGTH: 35895
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(35895)
; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1
US-10-767-471-10749

```

Query Match 9.3%; Score 30.2; DB 6; Length 35895;

201 GACTCTAACCTTTGGTCTCATTTGCCGTTGCCATTACAGTGGTGGCGCATCTTCTTGGCG 260

Db 12662 GCCTCTACCCACTGCTCTTGGTTGTCTTGCGACGACAGGCCCATCCTCCTGGCTG 12603

QY	261	TCGACGACCTGTTGAGGTTGGCGGCTGGACCATCGTTTGTGTGAGACGGGAATCT	320
Db	12602	CCACACATCTGCTTGCCCTCCTGACCTCCACGACACTGAGTAATGTTGACACTCAGACCT	126
QY	321	TTT	323
Db	12542	TCT	12540

```

RESULT 4
US-10-767-701-5423/C
; Sequence 5423, Application US/10767701
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ. ID NOS: 63128
; SEQ. ID NO 5423
; LENGTH: 731
; TYPE: DNA
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-CLUS32149_1
US-10-767-701-5423

```

Query Match	8.9%;	Score 28.8;	DB 6;	Length 731;
Best Local Similarity	53.6%;	Pred. No. 4.3;		
Matches 60;	Conservative 0;	Mismatches 52;	Indels 0;	Gaps 0;

149 TCCCTGCCATTTTGACCGCATCAACGTTTCGCAGCAATGCCATAGATCTGAAGACTTAA 208

Db 240 TTCTGCCAGTTTGTCTCTCGGGACACATCCACCATTGAATTGGTTCAGAGTAACATG 181

209 CCTTTGGTCTCATTCGCCGTTGCATTAAGTGTGTGGCGCATCTTTGGCGG 260

Db 180 CATCTCATCTTGAGTTCCGAGTTAGGAGCTGAACCAACTCTCGGGCTG 129

```

RESULT 5
US-10-767-471-10669/c
; Sequence 10669, Application US/10767471
; GENERAL INFORMATION.
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: RHEUMATOID ARTHRITIS, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO001505
; CURRENT APPLICATION NUMBER: US/10/767,471
; CURRENT FILING DATE: 2004-01-30
; NUMBER OF SEQ ID NOS: 50231
; SOFTWARE: FASTSEQ For Windows Version 4.0
; SEQ ID NO 10669
; LENGTH: 158027
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...T(158027)
; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1
US-10-767-471-10669

```

Query Match	8.9%	Score 28.8;	DB 6;	Length 158027;
Best Local Similarity	54.0%	Pred. No. 50;		
Matches	54;	Mismatches	44;	Indels 0;
		Conservative	2;	Gaps 0;

48 AGTCATTACTTTTGGCGCTCCGGCGGTTCCGTTCTTAATCCTTAAGCCCTACGTGATC 107

Db 48997 AGTCTTGAATCTTYACCCACAGCTGCTATGGTCTGAATATCTGTGTCCCTCCAAATTC 48938

Tue Apr 20 06:47:20 2004

us-10-608-504-4.rmpn

Page 1

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 17, 2004, 06:48:18 ; Search time 37.4407 Seconds
(without alignments)
5694.647 Million cell updates/sec

Title: US-10-608-504-4
Perfect score: 324
Sequence: 1 atgacactgattctctctc.....ttgactgtgactctttc 324

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 1362413 seqs, 329030602 residues

Total number of hits satisfying chosen parameters: 2724826

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database:

Pending Patents NA New:*
1: /cgn2_6/prodata/1/pna/US06_NEW_COMB_seq:*
2: /cgn2_6/prodata/1/pna/US07_NEW_COMB_seq:*
3: /cgn2_6/prodata/1/pna/US08_NEW_COMB_seq:*
4: /cgn2_6/prodata/1/pna/US09_NEW_COMB_seq:*
5: /cgn2_6/prodata/1/pna/US10_NEW_COMB_seq:*
6: /cgn2_6/prodata/1/pna/US10_NEW_COMB_seq:*
7: /cgn2_6/prodata/1/pna/US60_NEW_COMB_seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	33.2	10.2	173233	6 US-10-021-698A-716	Sequence 716, App
2	30.4	9.4	13350	6 US-10-789-400-36	Sequence 36, Appl
3	30.2	9.3	35895	6 US-10-767-471-10749	Sequence 10749, A
4	28.8	8.9	731	6 US-10-767-701-5423	Sequence 5423, Ap
5	28.8	8.9	158027	6 US-10-767-471-10669	Sequence 10669, A
6	28.6	8.8	2737	6 US-10-650-609-1	Sequence 1, Appl
7	28.6	8.8	3042	6 US-10-451-467A-545	Sequence 545, App
8	28.4	8.8	312	6 US-10-767-701-21673	Sequence 21673, A
9	28.4	8.8	500	6 US-10-108-605A-248	Sequence 767, App
10	28.4	8.8	8577	6 US-10-108-605A-248	Sequence 248, App
11	28.4	8.8	20618	6 US-10-796-280-12417	Sequence 12417, A
12	28.4	8.8	46075	6 US-10-796-280-12418	Sequence 12418, A
13	28.4	8.8	69359	6 US-10-796-280-12452	Sequence 12452, A
14	28.4	8.8	234309	6 US-10-796-280-12378	Sequence 12378, A
15	28.2	8.7	490	6 US-10-767-701-27384	Sequence 27384, A
16	28.2	8.7	531	6 US-10-767-701-25301	Sequence 25301, A
17	28.2	8.7	77294	6 PCT-US03-36500-1	Sequence 1, Appl
18	28.2	8.7	185248	6 US-10-796-280-12228	Sequence 12228, A
19	28	8.6	528	6 US-10-767-701-2761	Sequence 2761, Ap
20	28	8.6	593	6 US-10-767-701-29783	Sequence 29783, A
21	28	8.6	1102	6 US-10-021-698A-4460	Sequence 4460, Ap
22	28	8.6	1758	6 US-10-100-683-1035	Sequence 1035, Ap
23	28	8.6	2178	6 PCT-US04-05654-2655	Sequence 2655, Ap
24	28	8.6	4990	6 US-10-100-683-4553	Sequence 4553, Ap
25	27.8	8.6	1226	6 PCT-US04-09299-18	Sequence 18, Appl
26	27.8	8.6	1675	6 US-10-152-372-135	Sequence 135, App

27	27.8	8.6	1709	6 US-10-100-683-3274	Sequence 3274, Ap
28	27.6	8.5	201	6 US-10-767-471-27293	Sequence 27293, A
29	27.6	8.5	201	6 US-10-767-471-27441	Sequence 27441, A
30	27.6	8.5	201	6 US-10-796-307-18171	Sequence 18171, A
31	27.6	8.5	410	6 US-10-767-701-29976	Sequence 29976, A
32	27.6	8.5	35884	6 US-10-796-280-12303	Sequence 12303, A
33	27.6	8.5	37689	6 US-10-796-280-12200	Sequence 12200, A
34	27.6	8.5	41706	6 US-10-796-307-8720	Sequence 8720, Ap
35	27.4	8.5	509	6 US-10-767-701-18755	Sequence 18755, A
36	27.4	8.5	699	6 US-10-767-701-12271	Sequence 12271, A
37	27.4	8.5	1197	6 US-10-771-241-144	Sequence 144, App
38	27.4	8.5	1657	6 US-10-100-683-5402	Sequence 5402, Ap
39	27.2	8.4	76829	6 US-10-765-790-27	Sequence 27, Appl
40	27.2	8.4	600	7 US-60-545-213-175	Sequence 175, App
41	27.2	8.4	600	7 US-60-545-213-175	Sequence 4447, Ap
42	27.2	8.4	600	7 US-60-545-213-4447	Sequence 4447, Ap
43	27.2	8.4	600	7 US-60-545-213-4448	Sequence 4448, Ap
44	27.2	8.4	1365	6 US-10-417-884A-1659	Sequence 1659, Ap
45	27.2	8.4	1500	1 PCT-US04-09323-40	Sequence 40, Appl

ALIGNMENTS

RESULT 1
US-10-021-698A-716
Sequence 716, Application US/10021698A
GENERAL INFORMATION:
APPLICANT: KEITH, TIM
APPLICANT: LITTLE, RANDALL
APPLICANT: VAN BERDREWEH, PAUL
APPLICANT: DUPUIS, JOSEF
APPLICANT: DEL MASTRO, RICHARD
APPLICANT: SIMON, JASON
APPLICANT: ALLEN, KRISTINA
APPLICANT: PANDIT, SONIL
TITLE OF INVENTION: NUCLEOTIDE AND AMINO ACID SEQUENCES RELATING TO
FILE REFERENCE: 2976-4044US1
CURRENT APPLICATION NUMBER: US/10/021,698A
PRIORITY FILING DATE: 2001-10-22
PRIORITY FILING DATE: 2000-06-14
PRIORITY FILING DATE: 2000-06-14
NUMBER OF SEQ ID NOS: 6160
SOFTWARE: PatentIn 2.1
SEQ ID NO 716
LENGTH: 173233
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: modified_base
LOCATION: (88987)..(89086)
OTHER INFORMATION: a, t, c or g
FEATURE:
NAME/KEY: modified_base
LOCATION: (167286)..(167385)
OTHER INFORMATION: a, t, c or g
US-10-021-698A-716
Query Match
Best Local Similarity 55.1%; Pred. No. 1.5;
Matches 65; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

QY 200 AGACCTTACCTTGTCTCTACCTGCGGATGATGAGTGGCGGATCTTCTGCG 259
DB 84488 AGAACTTACCTTGTCTCTACCTGCGGATGATGAGTGGCGGATCTTCTGCG 84547
QY 260 GTGACGACCTTGTCTCTACCTGCGGATGATGAGTGGCGGATCTTCTGCG 317
DB 84548 GTAACTTACCTTGTCTCTACCTGCGGATGATGAGTGGCGGATCTTCTGCG 84605

RESULT 2

Tue Apr 20 06:47:19 2004

us-10-608-504-4.rn1

Page 7

ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &
ADDRESS: No. 603115015
STREET: One Liberty Place 46th. Floor
CITY: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19103
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/476,900A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Beardell, Lori Y.
REGISTRATION NUMBER: 34,293
REFERENCE/DOCKET NUMBER: BYLR-0027
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-568-3100
TELEFAX: 215-568-3439
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 5110 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: nucleic acid
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-476-900A-4

Query Match 9.2%; Score 29.8; DB 3; Length 5110;
Best Local Similarity 58.4%; Pred. No. 3;
Matches 52; Conservative 0; Mismatches 37; Indels 0; Gaps 0;
QY 118 GGCAGAAATGCGCATGTGATGCCAGAGGATCCTTGCCATTTTGACCGCATCAACGTTT 177
DB 4500 GGCAGCTTGCGCATGTGATGCCATGCATGATGATAGCCCTTCAACATATCCACCAT 4441
QY 178 CGCAGCAATGCCATAGATCTGAAGACTT 206
DB 4440 CGGAAAAAGCGCAGAGAGAGGAGATT 4412

Search completed: April 19, 2004, 18:38:15
Job time: 28.2807 secs

REFERENCE/DOCKET NUMBER: BYLR-0003
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-568-3100
TELEFAX: 215-568-3439
INFORMATION FOR SEQ ID NO: 27:
SEQUENCE CHARACTERISTICS:
LENGTH: 4635 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
FEATURE:
NAME/KEY: CDS
LOCATION: 37..4533
US-08-404-531B-27

Query Match 9.2%; Score 29.8; DB 2; Length 4635;
Best Local Similarity 58.4%; Pred. No. 2.8;
Matches 52; Conservative 0; Mismatches 37; Indels 0; Gaps 0;
CY 118 GCGAAATGGCGATGTGATGCGCAGAGATCTTCCATTGACCGCATCAAGCTT 177
DB 4025 GCGAGCTGGCGATGTGATGCGCAGATGATGATGCGCCTTCAACATATCCACATT 3966
CY 178 CGCAGCATGCGATGATGATGCGCAGATCTT 206
DB 3965 CGGAAAAAGCGCAGAGAGAGAGAGATTT 3937

RESULT 11
US-08-476-900A-27/c
Sequence 27, Application US/08476900A
PATENT NO. 6031150
GENERAL INFORMATION:
APPLICANT: Joseph Bryan, Lydia Aguilar Bryan, Daniel Nelson
TITLE OF INVENTION: Sequence Encoding Mammalian Sulfonylurea Receptor
Patent No. 6031150
TITLE OF INVENTION: and Method of Detecting Persistent Hyperinsulinemic Hypoglycemia
TITLE OF INVENTION: Infancy
NUMBER OF SEQUENCES: 49
CORRESPONDENCE ADDRESS:
ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &
ADDRESSEE: No. 6031150 is
STREET: One Liberty Place 46th. Floor
CITY: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19103
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/476,900A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Beardell, Lori Y.
REGISTRATION NUMBER: 34,293
REFERENCE/DOCKET NUMBER: BYLR-0027
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-568-3100
TELEFAX: 215-568-3439
INFORMATION FOR SEQ ID NO: 27:
SEQUENCE CHARACTERISTICS:
LENGTH: 4635 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear

MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
FEATURE:
NAME/KEY: CDS
LOCATION: 37..4533
US-08-476-900A-27

Query Match 9.2%; Score 29.8; DB 3; Length 4635;
Best Local Similarity 58.4%; Pred. No. 2.8;
Matches 52; Conservative 0; Mismatches 37; Indels 0; Gaps 0;
CY 118 GCGAAATGGCGATGTGATGCGCAGAGATCTTCCATTGACCGCATCAAGCTT 177
DB 4025 GCGAGCTGGCGATGTGATGCGCAGATGATGATGCGCCTTCAACATATCCACATT 3966
CY 178 CGCAGCATGCGATGATGATGCGCAGATCTT 206
DB 3965 CGGAAAAAGCGCAGAGAGAGAGAGATTT 3937

RESULT 12
US-08-488-546A-27/c
Sequence 27, Application US/08488546A
PATENT NO. 6054313
GENERAL INFORMATION:
APPLICANT: Joseph Bryan, Lydia Aguilar Bryan, Daniel Nelson, Pamela
APPLICANT: Thomas, Gilbert Cole, and Robert Gagel
TITLE OF INVENTION: Sequence Encoding Mammalian Sulfonylurea Receptor
Patent No. 6054313
NUMBER OF SEQUENCES: 49
CORRESPONDENCE ADDRESS:
ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &
ADDRESSEE: No. 6054313 is
STREET: One Liberty Place 46th. Floor
CITY: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19103
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/488,546A
FILING DATE: 07-JUNE-1995
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/404,531
FILING DATE: 15-MARCH-1995
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Beardell, Lori Y.
REGISTRATION NUMBER: 34,293
REFERENCE/DOCKET NUMBER: BYLR-0026
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-568-3100
TELEFAX: 215-568-3439
INFORMATION FOR SEQ ID NO: 27:
SEQUENCE CHARACTERISTICS:
LENGTH: 4635 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
FEATURE:
NAME/KEY: CDS
LOCATION: 37..4533

Page 4

	Query Match	9.5%	Score 30.8	DB 4	Length 5829
	Best Local Similarity	57.1%	Pred. No. 1.4		
	Matches 56	Conservative 0	Mismatches 42	Indels 0	Gaps 0
QY	122	AAATGGGAGTGTGATGACGAGAAATTCCTTGCATTTGACCGCATCAACCTTTGCGA	181		
DB	671	ATATGGGAGGAAATATTCATCGATTCTTTCGACCTTTCTCGACATATTTAAAAGG	612		
QY	182	GCAATGGCATGATCTGAAAGACTCTAACCTTGTGCTC	219		
DB	611	AAAGATGCTAAGATATGAGAAATATGATCGCTTTTGTATC	574		

```

TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
FAX: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: PTZSPC-F15
US-08-232-463-14

Query Match          9.4%; Score 30.6; DB 1; Length 7218;
Best Match Similarity 1.0%; Pred.No.1.9;
Matches      3; Conservative 179; Mismatches 133; Indels    0; Gaps    0;

QY      10 GATTTCCTGTAATCTCTTGTTGTGGCAGTATGCACTCATCTTACCTTGGCGTCGG 69
Db       |||.....: : : : : : : : : : : : : : : : : : : : : : : : : : : :
        1064 GATTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 1123
QY       70 GGCGTCGCTTCTATCCTTAAAGCCCTAAGTGAATACAATCTTGTGGGCAAAATGCG 129
Db       : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      1124 YTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 1188
Db       : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      130 ATGTGATGCCAGCAGAACTTGCATTTTGAACGCAATCAAGTTTGGCAGCAATGCG 189
Db       : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
        1184 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1243
QY      190 ATTAGTGAACACTCTAACCTTGGTCTCATGTCGCTTGCATTCAGTGGGCGGCGCAT 249
Db       : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
        1244 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1303
QY      250 CTCTGCGGCGTCAAGCACCTCTGTGAGCGTTGGCGCTGGCAGCAATCGTTTGTGGA 309
Db       : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
        1304 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1363
QY      310 CTGGTGATCTTTTC 324
Db       : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
        1364 YYYYYYYYYYYYYYY 1378

RESULT 10
US-08-404-531B-27/C
; Sequence 27, Application US/08404531B
; Patent No. 5863724
GENERAL INFORMATION:
APPLICANT: Joseph Bryan, Lydia Aguilar Bryan, Daniel Nelson, Pamela
            Applicant, Thomas, Gilbert Core, and Robert Gagel
TITLE OF INVENTION: Sequence Encoding Mammalian Sulfomylurea Receptor
Patent No. 5863724
NUMBER OF SEQUENCES: 49
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &
ADDRESSEE: No. 5863724ris
STREET: One Liberty Place 46th. Floor
CITY: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19103
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/404,531B
FILING DATE: 15-MAR-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Beardsell, Lori Y.
REGISTRATION NUMBER: 34,293

```

PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 13729
LENGTH: 594
TYPE: DNA
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-13729

Query Match
Best Local Similarity 10.6%; Score 34.2; DB 4; Length 594;
Best Local Similarity 56.8%; Pred. No. 0.023;
Matches 63; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

QY 204 TCTAACCTTGGTTCATTCGCGTGGCGATTAAGTGGTGGCGGATCTTCTGGCGGTCG 263
DB 407 TATAGGCGCTTGGCGCTTCGCGTGGCGGATTAAGTGGTGGCGGCGGCTTATCTTCTGATCC 348
QY 264 ACGCACCTTGTGAGCGTGGCGGCGTGGCGACCATCGTTTGTGGACTGGT 314
DB 347 GCGCGACGGGCGTTCGATCGGGGTGTGTCGGCGGCTTCTCGGTGGCCAGGT 297

RESULT 5
US-09-252-991A-13832/C
Sequence 13832, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubinfeld et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 13832
LENGTH: 678
TYPE: DNA
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-13832

Query Match
Best Local Similarity 10.6%; Score 34.2; DB 4; Length 678;
Best Local Similarity 56.8%; Pred. No. 0.023;
Matches 63; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

QY 204 TCTAACCTTGGTTCATTCGCGTGGCGATTAAGTGGTGGCGGATCTTCTGGCGGTCG 263
DB 472 TATAGGCGCTTGGCGCTTCGCGTGGCGGATTAAGTGGTGGCGGCGGCTTATCTTCTGATCC 413
QY 264 ACGCACCTTGTGAGCGTGGCGGCGTGGCGACCATCGTTTGTGGACTGGT 314
DB 412 GCGCGACGGGCGTTCGATCGGGGTGTGTCGGCGGCTTCTCGGTGGCCAGGT 362

RESULT 6
US-09-328-352-1339
Sequence 1339, Application US/09328352
Patent No. 6562958
GENERAL INFORMATION:
APPLICANT: Gary L. Breton et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
FILE REFERENCE: GTC99-039A
CURRENT APPLICATION NUMBER: US/09/328,352
CURRENT FILING DATE: 1999-06-04
NUMBER OF SEQ ID NOS: 8252
SEQ ID NO 1339
LENGTH: 1404
TYPE: DNA
ORGANISM: Acinetobacter baumannii

US-09-328-352-1339

Query Match
Best Local Similarity 9.8%; Score 31.6; DB 4; Length 1404;
Best Local Similarity 55.5%; Pred. No. 0.32;
Matches 61; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

QY 136 ATCCGACGAGAAATCTTGCATTTGACCGCATCAAGCTTTCGACGAATGCGATGAT 195
DB 541 ATTCGATGGAGATATGGGCAATTAATGACATTAAGCTATGCGCAATGTAAGAG 600
QY 196 CTGAAGACTTAACTTTGCTCATTCGCGTGGCGGATTAAGTGGTGGCG 245
DB 601 CCTACCGTTAAACTTTGATCTTATTTGATTTGATTTGCTCTGATGCTGCG 650

RESULT 7
US-08-961-527-56
Sequence 56, Application US/08961527
Patent No. 6420135
GENERAL INFORMATION:
APPLICANT: Charles Kunsch
TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
CORRESPONDENCE ADDRESS:
ADDRESS: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4MB storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/961,527
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Brookes, A. Anders
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PB340P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 56:
SEQUENCE CHARACTERISTICS:
LENGTH: 2387 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-961-527-56

Query Match
Best Local Similarity 9.7%; Score 31.4; DB 4; Length 2387;
Best Local Similarity 52.7%; Pred. No. 0.51;
Matches 68; Conservative 0; Mismatches 61; Indels 0; Gaps 0;

QY 112 TTGTGGGCAAAATGGGATGTGATGCCAGACGAATCTTGCAATTTTGACCGCATCA 171
DB 962 TTGTGGGCAAAATGGGATGTGATTTTGAAGAAAGCAATTAATAATTAATGCTTTA 1021
QY 172 ACGTTGGCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 231
DB 1022 GCAATTAAGCGTATTTCCAAATGAGAAATTAACAGATGATGCTTCTGCTTAAT 1061
QY 232 ATTACAGTG 240
DB 1082 ATGATTTG 1090

QY 241 GTGGCGCATCTTCTTGGCGGTGAGCGACCTTGTAGCGTTGGCGGTGACCATCGTT 300
DB 241 GTGGCGCATCTTCTTGGCGGTGAGCGACCTTGTAGCGTTGGCGGTGACCATCGTT 300
QY 301 TTGTGTGACTGTGTAATCTTTTC 324
DB 301 TTGTGTGACTGTGTAATCTTTTC 324

RESULT 2

US-09-471-803A-1
; Sequence 1, Application US/09471803A
; Patent No. 6613545
; GENERAL INFORMATION:
; APPLICANT: KENNERKNECHT, NICOLE
; APPLICANT: SAHM, HERMANN
; APPLICANT: EGGELING, LOTMAR
; APPLICANT: PFEFFERLE, WALTER
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCE CODING FOR THE EXPORT OF
; TITLE OF INVENTION: BRANCHED CHAIN AMINO ACIDS, PROCESS FOR THE
; TITLE OF INVENTION: ISOLATION THEREOF AND USE THEREOF
; FILE REFERENCE: 21123/265496/MAS
; CURRENT APPLICATION NUMBER: US/09/471,803A
; PRIOR FILING DATE: 1999-12-23
; PRIOR FILING DATE: 1999-10-27
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1271
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
; FEATURE:
; NAME/KEY: gene
; LOCATION: (101)..(853)
; OTHER INFORMATION: brnf
; NAME/KEY: gene
; LOCATION: (853)..(1176)
; OTHER INFORMATION: brnf
; OTHER INFORMATION: ATCC14752
US-09-471-803A-1

Query Match 100.0%; Score 324; DB 4; Length 1271;
Best Local Similarity 100.0%; Pred. No. 5e-107;
Matches 324; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGACACTGATTTCTCTGATATTCCTGTTGTCGAGATGTCAGTCACTTACTTT 60
DB 853 ATGACACTGATTTCTCTGATATTCCTGTTGTCGAGATGTCAGTCACTTACTTT 912
QY 61 GCGCTCGGGGGGTTCCGTTCTTAATCTTAAGCCCTTAAGCGTGAATCAATTTGGGGC 120
DB 913 GCGCTCGGGGGGTTCCGTTCTTAATCTTAAGCCCTTAAGCGTGAATCAATTTGGGGC 972
QY 121 AAATGGCGATGTGATGTCAGAGAGATCTTGCAATTTGACCGCATCAAGTTCCG 180
DB 973 AAATGGCGATGTGATGTCAGAGAGATCTTGCAATTTGACCGCATCAAGTTCCG 1032
QY 181 AGCAATGCGATGATCTGAAGACTTAACTTTGTTGTTCTCAATTCGCTTGCAATTAAGTG 240
DB 1033 AGCAATGCGATGATCTGAAGACTTAACTTTGTTGTTCTCAATTCGCTTGCAATTAAGTG 1092
QY 241 GTGGCGCATCTTCTTGGCGGTGAGCGACCTTGTAGCGTTGGCGGTGACCATCGTT 300
DB 1093 GTGGCGCATCTTCTTGGCGGTGAGCGACCTTGTAGCGTTGGCGGTGACCATCGTT 1152
QY 301 TTGTGTGACTGTGTAATCTTTTC 324
DB 1153 TTGTGTGACTGTGTAATCTTTTC 1176

RESULT 3
US-09-471-803A-6

; Sequence 6, Application US/09471803A
; Patent No. 6613545
; GENERAL INFORMATION:
; APPLICANT: KENNERKNECHT, NICOLE
; APPLICANT: SAHM, HERMANN
; APPLICANT: EGGELING, LOTMAR
; APPLICANT: PFEFFERLE, WALTER
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCE CODING FOR THE EXPORT OF
; TITLE OF INVENTION: BRANCHED CHAIN AMINO ACIDS, PROCESS FOR THE
; TITLE OF INVENTION: ISOLATION THEREOF AND USE THEREOF
; FILE REFERENCE: 21123/265496/MAS
; CURRENT APPLICATION NUMBER: US/09/471,803A
; PRIOR FILING DATE: 1999-12-23
; PRIOR FILING DATE: 1999-10-27
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 1271
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
; FEATURE:
; NAME/KEY: gene
; LOCATION: (101)..(853)
; OTHER INFORMATION: brnf
; NAME/KEY: gene
; LOCATION: (853)..(1176)
; OTHER INFORMATION: brnf
; OTHER INFORMATION: ATCC13032
US-09-471-803A-6

Query Match 100.0%; Score 324; DB 4; Length 1271;
Best Local Similarity 100.0%; Pred. No. 5e-107;
Matches 324; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGACACTGATTTCTCTGATATTCCTGTTGTCGAGATGTCAGTCACTTACTTT 60
DB 853 ATGACACTGATTTCTCTGATATTCCTGTTGTCGAGATGTCAGTCACTTACTTT 912
QY 61 GCGCTCGGGGGGTTCCGTTCTTAATCTTAAGCCCTTAAGCGTGAATCAATTTGGGGC 120
DB 913 GCGCTCGGGGGGTTCCGTTCTTAATCTTAAGCCCTTAAGCGTGAATCAATTTGGGGC 972
QY 121 AAATGGCGATGTGATGTCAGAGAGATCTTGCAATTTGACCGCATCAAGTTCCG 180
DB 973 AAATGGCGATGTGATGTCAGAGAGATCTTGCAATTTGACCGCATCAAGTTCCG 1032
QY 181 AGCAATGCGATGATCTGAAGACTTAACTTTGTTGTTCTCAATTCGCTTGCAATTAAGTG 240
DB 1033 AGCAATGCGATGATCTGAAGACTTAACTTTGTTGTTCTCAATTCGCTTGCAATTAAGTG 1092
QY 241 GTGGCGCATCTTCTTGGCGGTGAGCGACCTTGTAGCGTTGGCGGTGACCATCGTT 300
DB 1093 GTGGCGCATCTTCTTGGCGGTGAGCGACCTTGTAGCGTTGGCGGTGACCATCGTT 1152
QY 301 TTGTGTGACTGTGTAATCTTTTC 324
DB 1153 TTGTGTGACTGTGTAATCTTTTC 1176

RESULT 4
US-09-252-991A-13729/c
; Sequence 13729, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Mats J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR FILING DATE: 1998-02-18

D5 181 AGCAATGCCATAGATCTGAGACTCTAACCTTTGGTCTCATTGCCGTTGCCATTACAGTG 2

CC The present invention relates to oligonucleotide clones originating in
 CC barley (*Hordeum vulgare*) which contain single nucleotide polymorphisms
 CC (SNP). The oligonucleotides may be used for analysis of SNPs among barley
 CC varieties, identification of particular varieties and genotype-phenotype
 CC analysis, isolation of specific genes and creation of new varieties by
 CC transformation of barley varieties with them and production of new barley
 CC varieties with desired properties. The present sequence represents an
 CC oligonucleotide clone DNA sequence featured in the specification. The
 CC sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at <http://wipo.int/pub/published-pct-sequences>. (Updated on 27-OCT-2003 to
 CC standardise OS field)
 XX

SQ Sequence 470 BP; 122 A; 131 C; 102 G; 106 T; 0 U; 9 Other;

Query Match 9.9%; Score 32; DB 8; Length 470;

Best Local Similarity 51.9%; Pred. No. 2.2; Mismatches 63; Indels 0; Gaps 0;

Matches 68; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

QY	181	AGCAATGCGATGATCTGAAGACTCTAACCTTTGGTCTCATTCGCCGTTGCGATTACAGTG	240
DB	384	ATCAAGNACAGGAGNAGATAGGAGACCTTTGAGNCCATCCAGTTGAGGATGCCAG	325
QY	241	GTTGCGCATCTTTGGCGGTGACCGCACTTTGAGGCTTGGCGCTGGCACCATCTT	300
DB	324	GGTCAACCACTTGGCCACCCGCTGGGCAAGTGTTCACCATTTGGCAAGGGCACCAAGCCT	265
QY	301	TTTGTGAGCT	311
DB	264	TGGGTGAGCCT	254

Search completed: April 19, 2004, 11:42:03
 Job time : 153.886 secs

CC The present invention relates to oligonucleotide clones originating in
 CC barley (*Hordeum vulgare*) which contain single nucleotide polymorphisms
 CC (SNP). The oligonucleotides may be used for analysis of SNPs among barley
 CC varieties, identification of particular varieties and genotype-phenotype
 CC analysis, isolation of specific genes and creation of new varieties by
 CC transformation of barley varieties with them and production of new barley
 CC varieties with desired properties. The present sequence represents an
 CC oligonucleotide clone DNA sequence featured in the specification. The
 CC sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published-pct-sequences. (Updated on 27-OCT-2003 to
 CC standardise OS field)

CC Sequence 376 BP; 97 A; 105 C; 82 G; 92 T; 0 U; 0 Other;

Query Match 9.9%; Score 32.2; DB 8; Length 376;

Best Local Similarity 53.6%; Pred. No. 1.7; Mismatches 0; Gaps 0;

Matches 67; Conservative 0; Indels 58; Indels 0; Gaps 0;

QY 187 GCGATGATCTGAAGACTTAACCTTGTGTCATTCGCCGTTGGCATTACAGTGTGGCG 246
 DB 375 GAGAGAGAGAGACATTAAGGGAACCTTTGAGACCATCCAGTTGAGATGCCAGGGTCAC 316
 QY 247 CATCTTCTTGGCGGTGACGACACCTTGTGAGCGTGGCGTGGACCATCGTTTGT 306
 DB 315 CAGTTTCCACCCGTCGTGGGCAACGTTTCAACCATTTGGCAAGGCAACCAAGCCTTGGGTG 256
 QY 307 GGAAT 311
 DB 255 AGCCT 251

RESULT 14
 ACL20868/c
 ID ACL20868 standard; DNA; 392 BP.

AC 27-OCT-2003 (revised)
 DT 17-OCT-2003 (first entry)

DE DNA clone originating in barley containing SNP encoding sequence #10850.

KM Barley; single nucleotide polymorphism; SNP; genotype-phenotype analysis;

KM Gene; ss.

OS *Hordeum vulgare*; var. (cul.Akashiniki).

PN WO2003057877-A1.

PD 17-JUL-2003.

PF 16-DEC-2002; 2002WO-IB005403.

PR 20-DEC-2001; 2001JP-00387059.

PR 20-DEC-2001; 2001JP-00387131.

PR 20-DEC-2001; 2001JP-00403299.

PR 27-SEP-2002; 2002JP-00327515.

PA (UYN1-) UNIV JAPAN OKAYAMA.

PI Sato K, Takeda K, Kohara Y;

WPI; 2003-587127/55.

PT Single nucleotide polymorphism sites in barley varieties and DNA
 PT sequences containing them for analysis and identification of barley
 PT varieties and production of barley transformants with desired
 PT characteristics.

PS Disclosure; SEQ ID XX; 284pp; Japanese.

XX

CC The present invention relates to oligonucleotide clones originating in
 CC barley (*Hordeum vulgare*) which contain single nucleotide polymorphisms
 CC (SNP). The oligonucleotides may be used for analysis of SNPs among barley
 CC varieties, identification of particular varieties and genotype-phenotype
 CC analysis, isolation of specific genes and creation of new varieties by
 CC transformation of barley varieties with them and production of new barley
 CC varieties with desired properties. The present sequence represents an
 CC oligonucleotide clone DNA sequence featured in the specification. The
 CC sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published-pct-sequences. (Updated on 27-OCT-2003 to
 CC standardise OS field)

CC Sequence 392 BP; 100 A; 105 C; 85 G; 102 T; 0 U; 0 Other;

Query Match 9.9%; Score 32.2; DB 8; Length 392;

Best Local Similarity 53.6%; Pred. No. 1.7; Mismatches 0; Gaps 0;

Matches 67; Conservative 0; Indels 58; Indels 0; Gaps 0;

QY 187 GCGATGATCTGAAGACTTAACCTTGTGTCATTCGCCGTTGGCATTACAGTGTGGCG 246
 DB 391 GAGAGAGAGAGACATTAAGGGAACCTTTGAGACCATCCAGTTGAGATGCCAGGGTCAC 332
 QY 247 CATCTTCTTGGCGGTGACGACACCTTGTGAGCGTGGCGTGGACCATCGTTTGT 306
 DB 331 CAGTTTCCACCCGTCGTGGGCAACGTTTCAACCATTTGGCAAGGCAACCAAGCCTTGGGTG 272
 QY 307 GGAAT 311
 DB 271 AGCCT 267

RESULT 15
 ACL20918/c
 ID ACL20918 standard; DNA; 470 BP.

AC 27-OCT-2003 (revised)

DT 17-OCT-2003 (first entry)

DE DNA clone originating in barley containing SNP encoding sequence #10909.

KM Barley; single nucleotide polymorphism; SNP; genotype-phenotype analysis;

KM Gene; ss.

OS *Hordeum vulgare*; var. (cul.Haruna Nijo).

PN WO2003057877-A1.

PD 17-JUL-2003.

PF 16-DEC-2002; 2002WO-IB005403.

PR 20-DEC-2001; 2001JP-00387059.

PR 20-DEC-2001; 2001JP-00387131.

PR 20-DEC-2001; 2001JP-00403299.

PR 27-SEP-2002; 2002JP-00327515.

PA (UYN1-) UNIV JAPAN OKAYAMA.

PI Sato K, Takeda K, Kohara Y;

WPI; 2003-587127/55.

PT Single nucleotide polymorphism sites in barley varieties and DNA
 PT sequences containing them for analysis and identification of barley
 PT varieties and production of barley transformants with desired
 PT characteristics.

PS Disclosure; SEQ ID XX; 284pp; Japanese.

XX

Query Match 10.0%; Score 32.4; DB 6; Length 6672;
 Best Local Similarity 58.6%; Pred. No. 5.4;
 Matches 75; Conservative 0; Mismatches 51; Indels 2; Gaps 1;

QY 93 GCCCCTACGTGATCAATTTGTGGGCAAAATGGCGATGTGAGTCCAGAGGAATCT 152
 DB 5382 GCACACACGTGATGTGCACATGTGTGAAAGAGGGTGGGTAGACATGCAATCT 5323
 QY 153 TGGCAATTTGACCGCATCAAGCTTTCGACGATGCGATAGATCTGAAGCTTAACCT 212
 DB 5322 TGTCACTGCGC--CAACAGATATGCGAGAAAGTGGGTGGGTGTGTGATCACTG 5265

QY 213 TGGTCTCA 220
 DB 5264 TATCCCA 5257

RESULT 12
 ACF39600/C
 ID ACF39600 standard; cDNA; 6672 BP.
 XX ACF39600;
 XX 29-SEP-2003 (first entry)
 DE Human MHC class II transactivator encoding cDNA SEQ ID NO:3.
 XX Human; major histocompatibility complex class II transactivator;
 KM MHC class II transactivator; antisense modulation; immunosuppressive;
 KM antitubercular; antidiabetic; antineoplastic; antitumor; cytostatic;
 KM neurotrophic; neuroprotective; immunostimulant; autoimmune disorder;
 KM MHC class II transactivator inhibitor; infection; transplant rejection;
 KM diabetes; rheumatoid arthritis; cancer; Alzheimer's disease;
 KM multiple sclerosis; severe combined immunodeficiency disease; gene; ss.
 XX Homo sapiens.
 OS
 XX
 FH Key Location/Qualifiers
 FT 139..3531
 FT CDS /*tag= a
 FT /product= "MHC class II transactivator"
 XX
 XX WO2003050247-A2.
 XX PD 19-JUN-2003.
 XX 04-DEC-2002; 2002WO-US038616.
 XX 05-DEC-2001; 2001US-00006366.
 XX (ISIS-) ISIS PHARM INC.
 XX PA Bennett FC, Dobie KM;
 XX PI
 XX WPI; 2003-577294/54.
 XX DR P-PSDB; ABR82108.
 XX PT New antisense oligonucleotides for modulating MHC class II transactivator
 PT gene expression, particularly useful for treating autoimmune disorders
 PT such as transplant rejection, Alzheimer's disease, or multiple sclerosis,
 PT or infection.
 XX
 XX Example 13, Page 93-99, 129pp; English.
 XX
 XX The present invention describes a compound (I) that is 8-50 nucleobases
 CC in length: (a) targets a nucleic acid molecule encoding major
 CC histocompatibility complex (MHC) class II transactivator, and
 CC specifically hybridizes with the nucleic acid encoding the MHC class II
 CC transactivator, and inhibits the expression of MHC class II
 CC transactivator; or (b) specifically hybridizes with at least an 8-
 CC nucleobase portion of an active site on a nucleic acid molecule encoding
 CC MHC class II transactivator. (I) has immunosuppressive, antimicrobial,
 CC antidiabetic, antineoplastic, antitumor, cytostatic, neurotrophic,

CC neuroprotective and immunostimulant activities, and can be used as an MHC
 CC class II transactivator inhibitor. The MHC class II transactivator
 CC antisense oligonucleotides can be used for treating an animal having a
 CC disease or condition associated with MHC class II transactivator, e.g.,
 CC autoimmune disorder or infection. The antisense oligonucleotides can be
 CC used for inhibiting the expression of MHC class II transactivator in
 CC cells or tissues. In particular, these diseases include transplant
 CC rejection, diabetes, rheumatoid arthritis, cancer, Alzheimer's disease,
 CC multiple sclerosis, or severe combined immunodeficiency disease. The
 CC antisense compounds are useful for diagnostics, prophylaxis, or as
 CC research reagents or kits. The present sequence encodes a human MHC class
 CC II transactivator, which is used in an example from the present invention
 XX

QY Sequence 6672 BP; 1392 A; 1954 C; 1868 G; 1458 T; 0 U; 0 Other;
 SO

Query Match 10.0%; Score 32.4; DB 8; Length 6672;
 Best Local Similarity 58.6%; Pred. No. 5.4;
 Matches 75; Conservative 0; Mismatches 51; Indels 2; Gaps 1;

QY 93 GCCCCTACGTGATCAATTTGTGGGCAAAATGGCGATGTGAGTCCAGAGGAATCT 152
 DB 5382 GCACACACGTGATGTGCACATGTGTGAAAGAGGGTGGGTAGACATGCAATCT 5323
 QY 153 TGGCAATTTGACCGCATCAAGCTTTCGACGATGCGATAGATCTGAAGCTTAACCT 212
 DB 5322 TGTCACTGCGC--CAACAGATATGCGAGAAAGTGGGTGGGTGTGTGATCACTG 5265

QY 213 TGGTCTCA 220
 DB 5264 TATCCCA 5257

RESULT 13
 ACT20925/C
 ID ACT20925 standard; DNA; 376 BP.
 XX ACT20925;
 AC
 XX 27-OCT-2003 (revised)
 DT 17-OCT-2003 (first entry)
 XX
 XX DNA clone originating in barley containing SNP encoding sequence #10916.
 DE
 XX Barley; single nucleotide polymorphism; SNP; genotype-phenotype analysis;
 KM gene; ss.
 KM
 XX Hordeum vulgare; var. (cul. Akashinikil).
 OS
 XX
 XX WO2003057877-A1.
 XX PD 17-JUL-2003.
 XX 16-DEC-2002; 2002WO-IB005403.
 XX PF
 XX 20-DEC-2001; 2001JP-00387059.
 XX PR 20-DEC-2001; 2001JP-00387131.
 XX PR 20-DEC-2001; 2001JP-00403239.
 XX PR 20-DEC-2001; 2001JP-00403300.
 XX PR 27-SEP-2002; 2002JP-00327515.
 XX (UNYI-) UNIV JAPAN OKAYAMA.
 XX PA
 XX PI
 XX Sato K, Takeda K, Kohara Y;
 XX WPI; 2003-587127/55.
 XX
 XX Single nucleotide polymorphism sites in barley varieties and DNA
 PT sequences containing them for analysis and identification of barley
 PT varieties and production of barley transformants with desired
 PT characteristics.
 XX
 XX Disclosure; SEQ ID XX; 284pp; Japanese.
 XX

CC capable of detecting 1000 or more genes from *Drosophila*. The invention is
 CC useful in developmental biology and in elucidating cell signaling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA
 CC sequences (AB101840-AB16175) and the encoded proteins (AB57737-
 CC AB572072). The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 3144 BP; 1056 A; 676 C; 680 G; 732 T; 0 U; 0 Other;
 Query Match 10.1%; Score 32.8; DB 4; Length 3144;
 Best Local Similarity 52.1%; Pred. No. 2.8;
 Matches 73; Conservative 0; Mismatches 67; Indels 0; Gaps 0;

QY 169 TCACGTTTCGACGCAATGCGATGATCTGAGACTCTACTTGGTCTGACCTG 228
 DB 1512 TCTAAATTTCTATATATATATATATATATATATATATATATATATATAT 1453
 QY 229 GCGATTACAGTGGTGGCGCATCTTCTGGCGGTGACGCACTTGTGAGCGTGGCGCT 288
 DB 1452 TCGTTTGGGTTGCTGCTTTCTTTTTCGTCGCGCAGCAATCGTTGCTGCGTGA 1393
 QY 289 GGCACCATCGTTTGTGTG 308
 DB 1392 CACAGCTTATGTTGTGTG 1373

RESULT 9
 AAS59515/C
 ID AAS59515 standard; DNA; 66788 BP.

XX AC AAS59515;
 XX DT 13-FEB-2002 (first entry)

DE Propionibacterium acnes immunogenic protein encoding DNA #10.

XX SAPHO syndrome; synovitis; acne; pustulosis; hyperostosis; osteomyelitis;
 XX uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
 XX inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
 XX dermatological; osteopathic; neuroprotectant; ds.

XX Propionibacterium acnes.

XX WO200181581-A2.

XX 01-NOV-2001.

XX 20-APR-2001; 2001WO-US012865.

XX 21-APR-2000; 2000US-0199047P.

XX 02-JUN-2000; 2000US-0208841P.

XX 07-JUL-2000; 2000US-0216747P.

XX (CORI-) CORIXA CORP.

XX Skeiky YAM, Persing DH, Mitcham JL, Wang SS, Bhatia A;
 XX L'Almeida J, Zhang Y, Jen S, Carter D;

XX WPI; 2001-616774/71.

XX Propionibacterium acnes polypeptides and nucleic acids useful for
 XX PT vaccinating against and diagnosing infections, especially useful for
 XX PT treating acne vulgaris.

XX Claim 1; SEQ ID NO 10; 1069bp; English.

XX Sequences AAS59506-AAS59804 represent DNA molecules encoding
 CC propionibacterium acnes immunogenic polypeptides. The proteins and their
 CC associated DNA sequences are used in the treatment, prevention and
 CC diagnosis of medical conditions caused by *P. acnes*. The disorders include

CC SAPHO syndrome (synovitis, acne, pustulosis, hyperostosis and
 CC osteomyelitis), uveitis and endophthalmitis. *P. acnes* is also involved in
 CC infections of bone, joints and the central nervous system, however it is
 CC particularly involved in the inflammatory lesions associated with acne
 CC vulgaris. A method for detecting the presence or absence of *P. acnes* in a
 CC patient comprises contacting a sample with a binding agent that binds to
 CC the proteins of the invention and determining the amount of bound protein
 CC in the sample. The polypeptides may be used as antigens in the production
 CC of antibodies specific for *P. acnes* proteins. These antibodies can be
 CC used to downregulate expression and activity of *P. acnes* polypeptides and
 CC therefore treat *P. acnes* infections. The antibodies may also be used as
 CC diagnostic agents for determining *P. acnes* presence, for example, by
 CC enzyme linked immunosorbent assay (ELISA). This sequence encodes the
 CC polypeptides shown in AAU1313-42019 and AAU67473. Note: The sequence
 CC data for this patent did not form part of the printed specification, but
 CC was obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 66788 BP; 13534 A; 21369 C; 19347 G; 12530 T; 0 U; 8 Other;

Query Match 10.1%; Score 32.6; DB 4; Length 66788;
 Best Local Similarity 52.6%; Pred. No. 13;
 Matches 71; Conservative 0; Mismatches 64; Indels 0; Gaps 0;

QY 154 GCCATTTTGACCGCATGACGTTTCGACGCAATGCGATGATCTGAGACTCTTAACCTT 213
 DB 40085 GTCAATTCGATCGCGGCGAGCAACCGCCGCGCTTCTGACCGGCTGCTACGCT 40026
 QY 214 GGTCTCATTTGCCGTTGGGATTAAGTGTGGCGCATCTTTTGGCGGTGACGCACTTG 273
 DB 40025 GGGCGTTTCGACCGTCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 39966
 QY 274 TTGAGCGTTGGCGCT 288
 DB 39965 CTCAGGTTACGCT 39951

RESULT 10
 ACF64444/C
 ID ACF64444 standard; DNA; 66788 BP.

XX ACF64444;

XX 17-OCT-2003 (first entry)

DE Propionibacterium acnes DNA contig sequence #10.

XX Acne vulgaris; antiheparin; dermatological; antibacterial;
 XX immunostimulant; immune response; vaccine; ds.

XX Propionibacterium acnes.

XX WO2003033515-A1.

XX 24-APR-2003.

XX 11-OCT-2002; 2002WO-US032727.

XX 15-OCT-2001; 2001US-00978825.

XX (CORI-) CORIXA CORP.

XX Mitcham JL, Skeiky YAM, Persing DH, Bhatia A, Maisonneuve JL;
 XX Zhang Y, Wang S, Jen S, Lodes MT, Benson DR, Jones R, Carter D;

XX Barth B, Vallée-Douglas U;

XX WPI; 2003-381789/36.

XX New Propionibacterium acnes polypeptides and polynucleotides encoding the
 PT polypeptide, useful for diagnosing, preventing or treating acne vulgaris,
 PT or for stimulating an immune response specific for a *P. acnes* protein.

XX Claim 1; SEQ ID NO 10; 1481bp; English.

CC coryneform bacterium. Coryneform bacteria are useful for producing amino
 CC acids, nucleic acids, vitamins, saccharides and organic acids,
 CC particularly L-lysine. The present sequence is a nucleic acid described
 CC in the exemplification of the invention. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from the European Patent Office

XX Sequence 349980 BP; 79703 A; 91547 C; 98381 G; 80349 T; 0 U; 0 Other;

Query Match 100.0%; Score 324; DB 5; Length 349980;

Best Local Similarity 100.0%; Pred. No. 2,1e-99;

Matches 324; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 ATGACACTGATTTCTCTCTGTTCTCTGTTGCGAGATGCGAGTACTTACTTT 60
DB 277581 ATGACACTGATTTCTCTCTGTTCTCTGTTGCGAGATGCGAGTACTTACTTT 277640
QY 61 GCGCTCCGCGCGGTCCTGTTCTTAACTTAAAGCCCTACGTAATCAATTTGTGGGC 120
DB 277641 GCGCTCCGCGCGGTCCTGTTCTTAACTTAAAGCCCTACGTAATCAATTTGTGGGC 277700
QY 121 AAAATGGCGATGTGATGCGACAGAAATCCTTGCATTGACCGCATCAACGTTGCG 180
DB 277701 AAAATGGCGATGTGATGCGACAGAAATCCTTGCATTGACCGCATCAACGTTGCG 277760
QY 181 AGCAATGCGATGATCTGAAAGACTTAACTTGTCTCATATGCCGTTGCGATTACAGTG 240
DB 277761 AGCAATGCGATGATCTGAAAGACTTAACTTGTCTCATATGCCGTTGCGATTACAGTG 277820
QY 241 GTGGCGCATCTTCTTGGCGCGTGAAGCGACCTTGTGAGCGTTGGCGCGCATCGTT 300
DB 277821 GTGGCGCATCTTCTTGGCGCGTGAAGCGACCTTGTGAGCGTTGGCGCGCATCGTT 277880
QY 301 TTTGTTGACTGCTGTAATCTTTTC 324
DB 277881 TTTGTTGACTGCTGTAATCTTTTC 277904

```

RESULT 7

ADA71938/c

ID ADA71938 standard; DNA; 2000 BP.

AC ADA71938;

DT 20-NOV-2003 (first entry)

XX Rice gene, SEQ ID 5263.

XX Plant; bacterial infection; fungal infection; viral infection; rice;

KM gene; ds.

XX Oryza sativa.

XX WO2003000898-A1.

XX 03-JAN-2003.

XX 22-JUN-2001; 2001WO-IB001105.

XX 22-JUN-2001; 2001WO-IB001105.

XX 22-JUN-2001; 2001WO-IB001105.

XX (SYGN) SYNGENTA PARTICIPATIONS AG.

XX Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;

XX Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;

XX WPI; 2003-175290/17.

XX Identifying at least one gene involved in plant resistance or response to

PT pathogenic infection for conferring resistance or tolerance to a plant to

PT bacterial, fungal or viral infection by determining or detecting plant

PT gene expression.

PS Claim 27; SEQ ID NO 5263; 899bp; English.

XX The present invention relates to a method (M1) for identifying genes
 CC involved in plant resistance or response to pathogenic infection. M1
 CC comprises identifying a gene whose expression is significantly altered in
 CC the incompatible interaction of plant gene expression relative to
 CC expression of the gene in an uninfected plant, in a mutant plant that
 CC does not express a gene associated with response to pathogenic infection,
 CC or in a corresponding incompatible or compatible interaction. (M1) is
 CC useful for conferring resistance to resistance or tolerance to a plant to
 CC bacterial, fungal or viral infection. The present sequence was used to
 CC illustrate the invention.

XX Sequence 2000 BP; 336 A; 265 C; 284 G; 363 T; 0 U; 752 Other;

Query Match 10.7%; Score 34.8; DB 7; Length 2000;

Best Local Similarity 12.9%; Pred. No. 0.47; Mismatches 97; Indels 0; Gaps 0;

Matches 30; Conservative 105; Mismatches 97; Indels 0; Gaps 0;

```

QY 4 ACACTGATTTCTCTGATTTCTCTGTTGTCGAGATGCGACTACTTCTTGGC 63
DB 901 AYAAMRRRTYKMSRBMWTMTKAMTMTCMCAKMYMATGATMMWRMYTTCY 842
QY 64 CTCCGCGCGGTCCTGTTCTTAACTTAAAGCCCTACGTAATCAATTTGTGGCGAA 123
DB 841 AMTCACKRYKAMTKMTWTWACMRATSRMRAMGMRKRYMKRAYMRMRCKAGM 782
QY 124 ATGGCGATGTGATGCGACAGAGATCCTTGCATTGACCGCATCAACGTTGCGAGC 183
DB 781 ARWMSRYRKMVKVATRYRWKAMTMMWSRMRKSYRMSGMRMSAMRYCSRMKC 722
QY 184 AATCCGATGATCTGAAGACTTAACCTTGTCTCATTTGCCGTTGCGATTA 235
DB 721 AKTKYASARWTKRKSRYRRRWKRGWTKYRFRYRWSCHWTRPWSKR 670

```

RESULT 8

ABL27130/c

ID ABL27130 standard; DNA; 3144 BP.

AC ABL27130;

DT 26-MAR-2002 (first entry)

XX Drosophila melanogaster genomic polynucleotide SEQ ID NO 32863.

XX Drosophila; developmental biology; cell signalling; insecticide;

KM pharmaceutical; gene; ds.

XX Drosophila melanogaster.

XX WO200171042-A2.

XX 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US009231.

XX 23-MAR-2001; 2000US-0191637P.

XX 11-JUL-2000; 2000US-00614150.

XX (PEKE) PE CORP NY.

XX Venter JC, Adams M, Li PWD, Myers EW;

XX WPI; 2001-656860/75.

XX New isolated nucleic acid detection reagent for detecting 1000 or more

PT genes from Drosophila and for elucidating cell signaling and cell-cell

PT interactions.

PS Claim 1; SEQ ID NO 32863; 21bp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent

Qy 301 TTGTTGAGCTGGTGAATCTTTTC 324
 Db 1153 TTGTTGAGCTGGTGAATCTTTTC 1176

RESULT 5

AAH21112
 ID AAH21112 standard; DNA; 1271 BP.

AC AAH21112;

DT 05-SEP-2001 (first entry)

DE C. glutamicum DNA encoding brnF and brnE.

XX L-amino acid production; brnF; brnE; branched-chain amino acid;
 XX Corynebacterium; leucine; isoleucine; valine; medicine;
 XX animal nutrition; ds.

OS Corynebacterium glutamicum.

XX Key Location/Qualifiers

FT CDS 101..856
 FT /*tag= a
 FT /product= "brnF"
 FT 853..1179
 FT /*tag= b
 FT /product= "brnE"

FT CDS

PN EP1096010-A1.

PD 02-MAY-2001.

PF 11-OCT-2000; 2000EP-00122057.

PR 27-OCT-1999; 99DE-01051708.

PA (DEGS) DEGUSSA AG.
 PA (KERJ) FORSCHUNGSZENTRUM JUELICH GMBH.

PI Kennerknecht N, Eggeling L, Sahm H, Pfeifferle W;

DR WPI; 2001-391595/42.

XX New export genes from coryneform bacteria, useful for increasing
 PT fermentative production of branched-chain amino acids.

XX Claim 4 (i); Page 17-18; 23pp; German.

XX This invention describes a novel isolated polynucleotide (I) containing
 CC at least one sequence that (i) is 70% identical with a sequence that
 CC encodes at least one of 251 amino acid (aa) (3) or 108 aa (5)
 CC polypeptides; (ii) encodes a polypeptide at least 70% identical with (3)
 CC or (5); (iii) is the complement of (i) or (ii), or (iv) contains at least
 CC 15 consecutive bases from (i)-(iii). The invention also describes (a)
 CC protein derived from sequences of 753 bp (2) or 324 bp (4); (b)
 CC coryneform microorganisms, especially Corynebacterium, transformed with
 CC one or more (I), where these are replicative DNA; (c) production of
 CC branched-chain L-aa by fermentation of coryneform bacteria in which the
 CC brnE and/or brnF genes (or equivalent sequences) are amplified,
 CC especially overexpressed; and (d) method for isolating the brnE and/or
 CC brnF genes. (I) is used for transformation of coryneform bacteria being
 CC used for fermentative production of branched-chain amino acids,
 CC specifically leucine, isoleucine and valine, which are useful in medicine
 CC and animal nutrition. (I) can also be used as source of primers and
 CC probes for isolation of related sequences. Transformation with (I)
 CC increases yield of branched-chain amino acids. This sequence encodes the
 CC Corynebacterium glutamicum ATCC 13032 brnF and brnE proteins described in
 CC the method of the invention

XX Sequence 1271 BP; 265 A; 328 C; 318 G; 360 T; 0 U; 0 Other;

Query Match 100.0%; Score 324; DB 4; Length 1271;
 Best Local Similarity 100.0%; Pred. No. 1.6e-100; Indels 0; Gaps 0;
 Matches 324; Conservative 0; Mismatches 0;

Qy 1 ATGACACTGATTTCTCTGTAATTCCTTGTGTCAGATATGCACTCAATTAATTTT 60
 Db 853 ATGACACTGATTTCTCTGTAATTCCTTGTGTCAGATATGCACTCAATTAATTTT 912
 Qy 61 GCGCTCCGGGCGGCTTCCGTTCTTAATCCCTTAAGCCCTTAAGCAATTTGAGGC 120
 Db 913 GCGCTCCGGGCGGCTTCCGTTCTTAATCCCTTAAGCCCTTAAGCAATTTGAGGC 972
 Qy 121 AAATGCGCATGTGATGTCAGAGAGAAATCCTTGCAATTTTGACCGCATCAAGTTGCG 180
 Db 973 AAATGCGCATGTGATGTCAGAGAGAAATCCTTGCAATTTTGACCGCATCAAGTTGCG 1032
 Qy 181 AGCAATGCGATGATCTGAAGACTTAACTTGTGCTCATTTGCGGTTGCGATTACAGTG 240
 Db 1033 AGCAATGCGATGATCTGAAGACTTAACTTGTGCTCATTTGCGGTTGCGATTACAGTG 1092
 Qy 241 GTGCGCATCTTCTTGCGGTCGACGACCTTGTGAGCGTTGCGCGTGGCAATCGTT 300
 Db 1093 GTGCGCATCTTCTTGCGGTCGACGACCTTGTGAGCGTTGCGCGTGGCAATCGTT 1152
 Qy 301 TTGTTGAGCTGGTGAATCTTTTC 324
 Db 1153 TTGTTGAGCTGGTGAATCTTTTC 1176

RESULT 6

AAH64966
 ID AAH64966 standard; DNA; 349980 BP.

AC AAH64966;

DT 26-SEP-2001 (first entry)

DE C glutamicum coding sequence fragment SEQ ID NO: 1.

XX Coryneform bacterium; amino acid synthesis; vitamin; saccharide;

XX organic acid synthesis; ds.

XX Corynebacterium glutamicum.

XX EPI106790-A2.

PD 20-JUN-2001.

PF 18-DEC-2000; 2000EP-00127688.

PR 16-DEC-1999; 99JP-00377484.

PR 07-APR-2000; 2000JP-00159162.

PR 03-AUG-2000; 2000JP-00280988.

PA (KYOW) KYOWA HAKKO KOSYO KK.

PI Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;

PI Tateishi N, Senoh A, Ikeda M, Ozaki A;

DR WPI; 2001-376931/40.

XX Novel polynucleotides derived from Coryneform bacteria, for identifying
 PT mutation point of a gene, measuring expression of a gene, analyzing
 PT expression profile or pattern of a gene and identifying homologous gene.

XX Claim 7; SEQ ID NO 1; 246pp + Sequence Listing; English.

XX The present invention provides a number of nucleotide and protein
 CC sequences from the Coryneform bacterium Corynebacterium glutamicum. These
 CC are useful for identifying the mutation point of a gene derived from a
 CC mutant of coryneform bacterium, measuring expression amount and analyzing
 CC the expression profile or expression pattern of a gene derived from
 CC Coryneform bacterium, and identifying a homologue of a gene derived from

Accession	Sequence	Score	DB	Length	Mismatches	Indels	Gaps
XX	13-JUN-2001; 2001IDE-01028510.						
PA	(DEGS) DEGUSSA AG.						
PI	Farwick M, Moeckel B, Pfeifferle W, Bathe B, Huthmacher K;						
PI	WPI; 2003-279970/28.						
XX	New nucleic acid array useful for monitoring mRNA expression of						
XX	Corynebacterium glutamicum during fermentation, comprising nucleic acid						
PT	from Corynebacterium glutamicum.						
PS	Claim 1; Page 647; 709pp; German.						
XX	This invention describes a novel nucleic acid array involving						
CC	Corynebacterium glutamicum polymucleotides. The arrays are used to						
CC	analyse C. glutamicum, particularly for monitoring a fermentation process						
CC	to determine expression levels of C. glutamicum cellular mRNA. Such						
CC	monitoring particularly differentiates between expression levels of						
CC	different strains of C. glutamicum and allows the adjustment of different						
CC	culture and fermentation conditions. ACD00010-ACD02188 represent C.						
CC	glutamicum derived polymucleotides described in the disclosure of the						
CC	invention						
SQ	Sequence 324 BP; 60 A; 79 C; 80 G; 105 T; 0 U; 0 Other;						
Query Match	100.0%; Score 324; DB 7; Length 324;						
Best Local Similarity	100.0%; Pred. No. 8.8e-101;						
Matches 324; Conservative	0; Mismatches 0; Indels 0; Gaps 0						
QY	1 ATGACACTGATTTCTCTGTAATTCCTGTTGTGTCAGATATGTCAGTCAATTAATTCTTT 60						
Db	1 ATGACACTGATTTCTCTGTAATTCCTGTAATTCCTGTTGTGTCAGATATGTCAGTCAATTAATTCTTT 60						
QY	61 GGGCCCGGGGGGTTCCGTTCTTAATCCCTTAAGCCCGTACGTGATCAATTTGGGCG 120						
Db	61 GGGCTCCGGGCGGTTCCGTTCTTAATCCCTTAAGCCCGTACGTGATCAATTTGGGCG 120						
QY	121 AAAATGCGCATGTGATGTCGACAGGAAATCCTTGCCATTGTAACGCGATCAACGTTTCCG 180						
Db	121 AAAATGCGCATGTGATGTCGACAGGAAATCCTTGCCATTGTAACGCGATCAACGTTTCCG 180						
QY	181 AGCAATGGATGATGATGCTGAAGACTCAACCTTGTGCTGATGCGCGTGCAGTTACAGTG 240						
Db	181 AGCAATGGATGATGATGCTGAAGACTCAACCTTGTGCTGATGCGCGTGCAGTTACAGTG 240						
QY	241 GTGGCGCATCTTCTTGGCGCGTGCAGCGACCTTGTGAGCGTTGGCGCTGGCACCATCGTT 300						
Db	241 GTGGCGCATCTTCTTGGCGCGTGCAGCGACCTTGTGAGCGTTGGCGCTGGCACCATCGTT 300						
QY	301 TTGTTGGACGTGTGAATCTTTTC 324						
Db	301 TTGTTGGACGTGTGAATCTTTTC 324						
RESULT 4							
AAH21109							
ID	AAH21109 standard; DNA; 1271 BP.						
XX	AAH21109;						
AC							
DT	05-SEP-2001 (first entry)						
XX							
DE	C. glutamicum DNA encoding brnF and brnE.						
XX							
KW	L-amino acid production; brnF; brnE; branched-chain amino acid;						
KW	corynebacterium; leucine; isoleucine; valine; medicine;						
KW	animal nutrition; ds.						
XX							
OS	Corynebacterium glutamicum.						
XX							
Key	Location/Qualifiers						
FT	CDS 101..856						

FT		/+tag= a
FT		/product= "brnf"
FT	CDS	853..1179
FT		/+tag= b
FT		/product= "brne"
FT		
PX		
PN		Bp1096010-AI.
PD		02-MAY-2001.
PF		11-OCT-2000; 2000EP-00122057.
PX		
PR		27-OCT-1999; 99DE-01051708.
PX		
PA		(DEGS) DEGUSSA AG.
PX		(KERU) FORSCHUNGSZENTRUM JUELICH GMBH.
PI		Kennerknecht N, Eggeling L, Sahn H, Pfeifferle W;
DR		WPI; 2001-391595/42.
DR		P-PEDB; AAB86247, AAB86248.
PT		New export genes from corynebacterium bacteria, useful for increasing
PT		fermentative production of branched-chain amino acids.
PX		
PS		Claim 4 (1); Page 13; 23pp; German.
CC		This invention describes a novel isolated polynucleotide (1) containing
CC		at least one sequence that (i) is 70% identical with a sequence that
CC		encodes at least one of 251 amino acid (aa) (3) or 108 aa (5)
CC		polypeptides; (ii) encodes a polypeptide at least 70% identical with (3)
CC		or (5); (iii) is the complement of (i) or (ii), or (iv) contains at least
CC		15 consecutive bases from (1)-(111). The invention also describes (a)
CC		corynebacterium microorganisms, especially Corynebacterium, transformed with
CC		one or more (1), where these are replicative DNA; (c) production of
CC		branched-chain L-a by fermentation of corynebacterium bacteria in which the
CC		brnE and/or brnF genes (or equivalent sequences) are amplified,
CC		especially overexpressed; and (d) method for isolating the brnE and/or
CC		brnF genes. (1) is used for transformation of corynebacterium bacteria being
CC		used for fermentative production of branched-chain amino acids,
CC		specifically leucine, isoleucine and valine, which are useful in medicine
CC		and animal nutrition. (1) can also be used as source of primers and
CC		probes for isolation of related sequences. Transformation with (i)
CC		increases yield of branched-chain amino acids. This sequence encodes the
CC		Corynebacterium glutamicum ATCC 14752 brnF and brnE proteins described in
CC		the method of the invention
XX		
SQ		Sequence 1271 BP; 265 A; 328 C; 318 G; 360 T; 0 U; 0 Other;
QY		
Query Match	100.0%; Score 324; DB 4; Length 1271;	
Best Local Similarity	100.0%; Pred. No. 1.6e-100;	
Matches 324; Conservative	0; Mismatches 0; Indels 0; Gaps 0;	
Dy	1 ATGACACTGATTTTCCTGTATTTCCTTGTGTGGAGTAGTGAGTCATTACTTTT 60	
Dy	853 AAGCAACAAGATTTCCTCGTATTCTCTTGTTGGCAGATATGGACGATTAACATT 912	
QY	61 GGCGTCGGGGCGCGTTCGTTCTTAATCCTTAAGCCCCCAAGCATGCATTTGTGGGC 120	
Dy	913 GGCGTCGGGGCGCGTTCGTTCTTAATCCTTAAGCCCCCAAGCATGCATTTGTGGGC 972	
QY	121 AAAATGGCATGTGATGGACGAGCAGAAATCCTTGCCATTGTGACCAGCATCAAGCTTGGC 180	
Dy	973 AAAATGGCATGTGATGGACGAGCAGAAATCCTTGCCATTGTGACCAGCATCAAGCTTGGC 1033	
QY	181 AGCAATGGCATATGATCTGAAGAAGCTTAACCTTTGGTCTTCATTTGCCGTTGCCATTACAGTG 240	
Dy	1033 AGCAATGGCATATGATCTGAAGAAGCTTAACCTTTGGTCTTCATTTGCCGTTGCCATTACAGTG 1093	
QY	241 GTGGGCGCATCTTCTGGGCGGTGAGCAGCAACTTGTGAGCGTGGCGGTGGACCATCGTT 300	
Dy	1093 GTGGGCGCATCTTCTGGGCGGTGAGCAGCAACTTGTGAGCGTGGCGGTGGACCATCGTT 1155	

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 19, 2004, 10:56:53 ; Search time 150.886 Seconds
(without alignments)
9122.247 Million cell updates/sec

Title: US-10-608-504-4

Sequence: 324
1 atgacaactgattctctctg.....ttgactggtgacattcttc 324

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: N_Geneseq_290404:*
2: geneseqn1980s:*
3: geneseqn1990s:*
4: geneseqn2000s:*
5: geneseqn2001as:*
6: geneseqn2001bs:*
7: geneseqn2003as:*
8: geneseqn2003bs:*
9: geneseqn2003cs:*
10: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being predicted, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	324	100.0	324	4 AAH21111	Aah21111 C. glutam
2	324	100.0	324	5 AAH65255	Aah65255 C. glutam
3	324	100.0	324	7 ACA01970	ACA01970 C. glutam
4	324	100.0	1271	4 AAH21109	Aah21109 C. glutam
5	324	100.0	1271	4 AAH21112	Aah21112 C. glutam
6	324	100.0	349980	5 AAH64966	Aah64966 C. glutam
7	34.8	10.1	2000	7 ADA71938	Ada71938 Rice gene
8	32.8	10.1	3144	4 ABL27130	AbL27130 Drosophila
9	32.6	10.1	66788	4 AAS59515	Aas59515 Propionib
10	32.6	10.1	66788	7 ACFF6444	Acff6444 Propionib
11	32.4	10.0	6672	6 ABF68121	Abf68121 Ovary can
12	32.4	10.0	6672	8 ACP38600	AcP38600 Human MHC
13	32.2	9.9	376	8 ACL20925	ACL20925 DNA clone
14	32.2	9.9	392	8 ACL20868	ACL20868 DNA clone
15	32	9.9	470	8 ACL20918	ACL20918 DNA clone
16	32	9.9	476	8 ACL20933	ACL20933 Human CDN
17	32	9.9	820	4 AAH05013	Aah05013 Human CDN
18	32	9.9	1808	4 AAH18459	Aah18459 Human CDN
19	31.8	9.8	2091	8 ADB06513	AdB06513 Alloiococ
20	31.8	9.8	2091	8 ADB06509	AdB06509 Alloiococ
21	31.8	9.8	2091	8 ADB06511	AdB06511 Alloiococ
22	31.8	9.8	110000	8 ADB12064_00	ADB12064 Alloiococ
23	31.8	9.8	110000	8 ADB12064_01	ADB12064 Alloiococ

24	31.6	9.8	1404	8 ADA30052	Ada30052 DNA encod
25	31.4	9.7	2387	2 AAH52189	Aah52189 Streptococ
26	31.4	9.7	110000	5 AAB56454_07	Continuation (8 of
27	31	9.6	4590	5 AAH24065	Aah24065 Yeast AOD
28	31	9.6	6415	2 AAH13212	Aah13212 Enterococ
29	31	9.6	6415	6 ABB99007	Abb99007 Enterococ
30	30.8	9.5	5949	2 AAT30053	Aat30053 S. pneumo
31	30.8	9.5	265118	5 AAH41227	Aah41227 Pyrococcu
32	30.6	9.4	413	8 ACL20863	ACL20863 DNA clone
33	30.6	9.4	789	7 ACA50939	ACA50939 Prokaryot
34	30.4	9.4	386	8 ACL20939	ACL20939 DNA clone
35	30.4	9.4	4739	6 ABX03452	Abx03452 Negative-
36	30.2	9.3	373	8 ACL20861	ACL20861 DNA clone
37	30.2	9.3	448	8 ACL20929	ACL20929 DNA clone
38	30.2	9.3	505	8 ACL20950	ACL20950 DNA clone
39	30.2	9.3	515	8 ACL20954	ACL20954 DNA clone
40	30.2	9.3	528	8 ACL20920	ACL20920 DNA clone
41	30.2	9.3	529	8 ACL20912	ACL20912 DNA clone
42	30.2	9.3	533	8 ACL20906	ACL20906 DNA clone
43	30.2	9.3	537	8 ACL20947	ACL20947 DNA clone
44	30.2	9.3	538	8 ACL20907	ACL20907 DNA clone
45	30.2	9.3	540	8 ACL20924	ACL20924 DNA clone

ALIGNMENTS

RESULT 1
AAH21111
ID AAH21111 standard; DNA; 324 BP.
XX
AC AAH21111;
XX
DT 05-SEP-2001 (first entry)
XX
XX C. glutamicum brne DNA.
DE
XX l-amino acid production; brnf, brne; branched-chain amino acid;
KW coryneform bacterium; leucine; isoleucine; valine; medicine;
KW animal nutrition; ds.
XX
OS Corynebacterium glutamicum.
XX
FN EPI096010-Al.
XX
PD 02-MAY-2001.
XX
PF 11-OCT-2000; 2000EP-00122057.
XX
PR 27-OCT-1999; 99DE-01051708.
XX
PA (DEGS) DEGUSA AG.
XX (KERJ) FORSCHUNGSZENTRUM JUELICH GMBH.
XX
DR Kennertknecht N, Eggeling L, Sahn H, Pfefferte W;
XX WPI, 2001-391595/42.
XX F-PsDB; AAB86248.
XX
PT New export genes from coryneform bacteria, useful for increasing
XX fermentative production of branched-chain amino acids.
XX
PS Claim 5; Page 16; 23pp; German.
XX
CC This invention describes a novel isolated polynucleotide (I) containing
CC at least one sequence that (i) is 70% identical with a sequence that
CC encodes at least one of 251 amino acid (aa) (3) or 108 aa (5)
CC polypeptides; (ii) encodes a polypeptide at least 70% identical with (3)
CC or (5); (iii) is the complement of (i) or (ii), or (iv) contains at least
CC 15 consecutive bases from (i) - (iii). The invention also describes (a)
CC protein derived from sequences of 753 bp (2) or 324 bp (4); (b)
CC coryneform microorganisms, especially Corynebacterium, transformed with
CC one or more (i), where these are replicative DNA; (c) production of


```

/product="putative mannosyltransferase A"
/protein_id="AA064208.1"
/db_xref="GI:28974207"
/translation="1HDILPITSPQFPISCELLQWIERIKRDPVLATSGTKN
DLKQEPKIKITIPGNSFTELKVRHIEKRTFLAVGTIERAKAKAIYVTFALQN
EGIEDLITVKGIGRVEDFEKELCKKHNIKYLKVSDFRLGSLYSTELIYASVD
EGYGLPMVEASIVGACITARDIPVFRDLIGENATYFDDISLVTIINIKYSSKINRI
GITWEYAREIILIKEDL"
gene
711..1856
/gene="mtfb"
711..1856
/gene="mtfb"
/feature="Cf0002"
/codon_start=1
/translation="1"
/product="putative mannosyltransferase B"
/protein_id="AA064209.1"
/db_xref="GI:28974208"
/translation="MINILIDTRPLASNLGTGRTYVEIAKLNNEKFPYDYGFIL
DELYGDNQOKKPSIKSHLINIAKRIIDKPTIKKVKQKLIKSLREKSLINDL
YQENRPIELKAKYTTVTHDLAPSYPERPERIDPEONFIPNLSTKDHITVS
NAIKRGIKLNISADKISVINGDENI FRKPKQTINIKLPELNNPILFVGGI
EPRKULITLQAVNELNINIDLVAGAGWENSISHLIQNNEHIFLGFPTDDLA
LYSSATIPFVPSIYEGFIPLEIAACAPILIDIEVFRPIYGVNVAEFSPNAKS
LAEKLNINNSDKLSIMKNGLLISKYSWKCIQEHKELFKILD"
2182..5010
/gene="sapA"
2182..5010
/gene="sapA"
/feature="Cf0003; similar to sap"
/codon_start=1
/translation="1"
/product="surface array protein"
/protein_id="AA064210.1"
/db_xref="GI:28974209"
/translation="MLNKTDVSMLYTTINGMASSGDNKRYMLDYANNNTLGVSLANI
MDSPGNAKFPDGLSLAGNEKDPVTKIYISALGNISDVGNVTKAITGGEPFDSK
GNIVSVASLSKGLIGAMINSMWNGSASKAIFPKAKAASVPADATLGGKISLDE
GYTSKLISEINSSASDIDKYSIEDALKSELPPNGSYDLEGNMLKGTDLDTFNGE
TYVNGTNSKSLSAFDKIDGAGRDITLNQVANNASAPAKISEDLVYALGVNVE
SNIVVDLPNDRTILNDSGTGMANLALDVGSKLAGLITDISKLNINIKNSVDALT
AANKSVNDVAGAVTLTDPKRAVNVKVAAGDVTLTDEASASISVDANGKVTLT
KATNVENLKYKATNLTITRGGKNTVTLBNVTLDDSSAAAIIDPKATLTANNVY
KIMDGSDSKLAHLKSSAETLNLNFSKEATFALVTVKANTLNADSNPFLITKE
DGDHDAALANDSFTFVYKNGKELFADGLDVTISIDSGENAKIVFGSNG
DALAGKVSAGANDLVLESNKLSAGSVIDGAGDITVLKSSALADATLGMKNI
EAVOVKLTANTNISTSAFPTIIVAGDETSOTKLINDDOKIDLAATLAAKSKITM
DLNEKSGDNDISILNNAKTADNNVTANGADKQIVIEGLSVITVSAVDSITANG
SIDKSPDUNSIILNNGAGITLKGFTVADYKXVSMIDASALKAAGLTPASAITANG
ATTKSGGADSTIVKGLANLTIGBSGDTYTLKGGTSIDVTIINDRSDDKIVLDGT
TFASAKAIEKTLANTLTGTSFSDVNOASGNGNTPIVKIFHYONDITYVVDKAGATP
TKADIVIKLAGIHELGTQNTITADSQ"
gene
complement (5749..9069)
/gene="sapA5"
complement (5749..9069)
/feature="Cf0004; similar to sap"
/codon_start=1
/translation="1"
/product="surface array protein"
/protein_id="AA064211.1"
/db_xref="GI:28974210"
/translation="MLNKTDVSMLYTTINGMASSGDNKRYMLDYANNNTLGVSLANI
MDSPGNAKFPDGLSLAGNEKDPVTKIYISALGNISDVGNVTKAITGGEPFDSK
GNIVSVASLSKGLIGAMINSMWNGSASKAIFPKAKAASVPADATLGGKISLDE
GYTSKLISEINSSASDIDKYSIEDALKSELPPNGSYDLEGNMLKGTDLDTFNGE
TYVNGTNSKSLSAFDKIDGAGRDITLNQVANNASAPAKISEDLVYALGVNVE
SNIVVDLPNDRTILNDSGTGMANLALDVGSKLAGLITDISKLNINIKNSVDALT
AANKSVNDVAGAVTLTDPKRAVNVKVAAGDVTLTDEASASISVDANGKVTLT
KATNVENLKYKATNLTITRGGKNTVTLBNVTLDDSSAAAIIDPKATLTANNVY
KIMDGSDSKLAHLKSSAETLNLNFSKEATFALVTVKANTLNADSNPFLITKE
DGDHDAALANDSFTFVYKNGKELFADGLDVTISIDSGENAKIVFGSNG
DALAGKVSAGANDLVLESNKLSAGSVIDGAGDITVLKSSALADATLGMKNI
EAVOVKLTANTNISTSAFPTIIVAGDETSOTKLINDDOKIDLAATLAAKSKITM
DLNEKSGDNDISILNNAKTADNNVTANGADKQIVIEGLSVITVSAVDSITANG
SIDKSPDUNSIILNNGAGITLKGFTVADYKXVSMIDASALKAAGLTPASAITANG
ATTKSGGADSTIVKGLANLTIGBSGDTYTLKGGTSIDVTIINDRSDDKIVLDGT
TFASAKAIEKTLANTLTGTSFSDVNOASGNGNTPIVKIFHYONDITYVVDKAGATP
TKADIVIKLAGIHELGTQNTITADSQ"
gene
complement (5749..9069)
/gene="sapA5"
complement (5749..9069)
/feature="Cf0004; similar to sap"
/codon_start=1
/translation="1"
/product="surface array protein"
/protein_id="AA064211.1"
/db_xref="GI:28974210"
/translation="MLNKTDVSMLYTTINGMASSGDNKRYMLDYANNNTLGVSLANI
MDSPGNAKFPDGLSLAGNEKDPVTKIYISALGNISDVGNVTKAITGGEPFDSK
GNIVSVASLSKGLIGAMINSMWNGSASKAIFPKAKAASVPADATLGGKISLDE
GYTSKLISEINSSASDIDKYSIEDALKSELPPNGSYDLEGNMLKGTDLDTFNGE
TYVNGTNSKSLSAFDKIDGAGRDITLNQVANNASAPAKISEDLVYALGVNVE
SNIVVDLPNDRTILNDSGTGMANLALDVGSKLAGLITDISKLNINIKNSVDALT
AANKSVNDVAGAVTLTDPKRAVNVKVAAGDVTLTDEASASISVDANGKVTLT
KATNVENLKYKATNLTITRGGKNTVTLBNVTLDDSSAAAIIDPKATLTANNVY
KIMDGSDSKLAHLKSSAETLNLNFSKEATFALVTVKANTLNADSNPFLITKE
DGDHDAALANDSFTFVYKNGKELFADGLDVTISIDSGENAKIVFGSNG
DALAGKVSAGANDLVLESNKLSAGSVIDGAGDITVLKSSALADATLGMKNI
EAVOVKLTANTNISTSAFPTIIVAGDETSOTKLINDDOKIDLAATLAAKSKITM
DLNEKSGDNDISILNNAKTADNNVTANGADKQIVIEGLSVITVSAVDSITANG
SIDKSPDUNSIILNNGAGITLKGFTVADYKXVSMIDASALKAAGLTPASAITANG
ATTKSGGADSTIVKGLANLTIGBSGDTYTLKGGTSIDVTIINDRSDDKIVLDGT
TFASAKAIEKTLANTLTGTSFSDVNOASGNGNTPIVKIFHYONDITYVVDKAGATP
TKADIVIKLAGIHELGTQNTITADSQ"

```

```

NLVNTGAFSIAATATSLKLNINLSINGVSADLNSVNVGTATLASLEANNIVSGEFLK
GTTTAKGDVDPNENVALTLGATITSGNSAVIISATGAVTIGAVATOGNLTINA
GNITLGNITIGAKGDIYVADLGGVGTGTINSNANVSITSNBVTVGSEISKNVET
AAAGGDLNLAQVITGGAAADALITIGKEDTQITTSASGLSGTTLTLTETATKSLD
ISAGVLSKATIDIDKNVSENLIVDIOGSPAETITANSISATLITISGLDGGG
ANTVAVAPDAAVATITIDLSGISAOTGLSGTITNNAQTLITTKSAGNDITTVG
INADITVAGAGNDVENVYTAKITVANTPERATITDPSAGSIFPAASVAYKSTV
DLSGRADLSAIAVLTDSDEATVYGYGTYNESLYYVAVTTTAAANDVLYLGTG
TVDLSLTVNNDIVEA"
complement (9374..10762)
/codon_start=1
/translation="1"
/evidence="not experimental"
/feature="Cf0005"
/product="hypothetical protein Cf0005"
/protein_id="AA064212.1"
/db_xref="GI:28974211"
/translation="MONIINONKMSSEPYKLYRQWPKLINSIKLRIHVLQGR
RSGKSTVFKLLINHLKSDSKELIYINIDDPFSTKSDKAVINDIYLSLQRLNKR
VKYFLDELSAVNWKRYKSAVDSLPKRTITSGNSALNTQVATLISGRYLSOI
YPLSFEIILKINNINDYFELIKENHVLVQDMAKFGSVFVTSNENKRDITISY
YDAILKDCVANNQIRDIAGFRELGFSLNFGALYSVNSLNAVKALENVQVETIGY
LQCYLDCDELKMFYSLEKQNNKKLYLDCNGFLSLAFSPDNKRALENLVFCELY
KLGFLFPYKUNDSKCDPLARKDKSIALQVCEINELKRIEIGLNTLPGVNEY
LTPDDRMETELSQFRLGLNLDVYKRPDGGDENSETREKSEKELKIMPFYEF
FSEFNETNLAL"
complement (10992..14765)
/gene="sapA6"
complement (10992..14765)
/gene="sapA6"
/feature="Cf0006; similar to sap"
/codon_start=1
/translation="1"
/product="surface array protein"
/protein_id="AA064213.1"
/db_xref="GI:28974212"

```

Query Match 11.8%; Score 38.2; DB 1; Length 65865;
 Best Local Similarity 47.0%; Pred. No. 2.4;
 Matches 118; Conservative 0; Mismatches 133; Indels 0; Gaps 0;

```

5 CAACGTATTTCTCCTGATTTCTCCTGTTGTCGAGTATGTCAGTATTACTTTGGCC 64
45075 CAATCCTGTTGCTGCTATTTGTTATTTTGGTCGATTTGCGAGTATTGTACGCG 45016
65 TCCGGCGGCTTCCTGCTTAATCCTTAAGCCCTACGAGTACAAATTGTCGGCAAAA 124
45015 TTTAGCTGCTCCGCTGTTTAAATTCACACTAGTAGCAGCTTTGCACTGGTTAATG 44956
125 TGGCAGTGTGATGCGACAGAAATCCTTGCATTTTGACCGCATCAAGCTTTGGCAGCA 184
44955 TGATAGTGTGTTGACAGATTAAATTACGCTTGTTGCGGAATCGGCCCTTAAGGGCAG 44896
185 ATGCATATGATCTAAGACTTAACCTTGGCTCATTCGCGGTTGGCATTAAGTGTGG 244
44895 AGATTTTACCGGCGCAACATGCACTTGCTGTTCCGCTGCGGTGAACAGATTATCG 44836
245 CGCATCTTCTT 255
44835 TTCCTGTTGT 44825

```

Search completed: April 19, 2004, 18:30:01
 Job time : 1377.24 secs

DEFINITION Sequence 1 from Patent EP1108790.
 ACCESSION AX120085 AX114121
 VERSION AX120085.1 GI:114036800
 KEYWORDS Corynebacterium glutamicum
 SOURCE Corynebacterium glutamicum
 ORGANISM Bacteria; Actinobacteriales; Actinomycetales;
 Corynebacteriaceae; Corynebacterium.
 REFERENCE 1 Nakagawa, S., Mizoguchi, H., Ando, S., Hayashi, M., Ochiai, K.,
 Yoko, H., Tateishi, N., Senoh, A., Ikeda, M. and Ozaki, A.
 TITLE Novel polynucleotides
 JOURNAL Patent: EP 1108790-A 1 20-JUN-2001;
 KYOWA HAKKO KOGYO CO., LTD. (JP)
 FEATURES
 source Location/Qualifiers
 1..349980 "Corynebacterium glutamicum"
 /organism="Corynebacterium glutamicum"
 /mol_type="genomic DNA"
 /db_xref="taxon:11718"
 /note="Seq 1 to long (3,309,400) split in 11, seq 1
 0.000.001 0.349.980 349.980"
 ORIGIN
 Query Match 100.0%; Score 324; DB 6; Length 349980;
 Best Local Similarity 100.0%; Pred. No. 1.1e-88;
 Matches 324; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 277581 ATGACACGATGTTCTTCCTGATATTCCTGTTGCGAGTATGTCAGTCACTTCTTT 60
 Qy 1 ATGACACGATGTTCTTCCTGATATTCCTGTTGCGAGTATGTCAGTCACTTCTTT 60
 Db 277581 ATGACACGATGTTCTTCCTGATATTCCTGTTGCGAGTATGTCAGTCACTTCTTT 277640
 Qy 61 GCGCTCCGGCGGCTTCCTGTTCTTAACTTAAAGCCCTTACGTAATCAATTTGTGGC 120
 Db 277641 GCGCTCCGGCGGCTTCCTGTTCTTAACTTAAAGCCCTTACGTAATCAATTTGTGGC 277700
 Qy 121 AAAATGGCGATGCGATGCGAGAGAACTTGGCACTTTTGACCGCATCAACGTTTGGC 180
 Db 277701 AAAATGGCGATGCGATGCGAGAGAACTTGGCACTTTTGACCGCATCAACGTTTGGC 277760
 Qy 181 AGCAATGCGATGATGATGAGACTTGAAGCTTGTCTCATTCGCGTTGCGATTACAGTG 240
 Db 277761 AGCAATGCGATGATGATGAGACTTGAAGCTTGTCTCATTCGCGTTGCGATTACAGTG 277820
 Qy 241 GTGGCGATTTCTTGTGGCGGTGACGACACTTTTGAAGCTTGGCGCTGACCACTGCT 300
 Db 277821 GTGGCGATTTCTTGTGGCGGTGACGACACTTTTGAAGCTTGGCGCTGACCACTGCT 277880
 Qy 301 TTGTTGACTGGTGAATCTTTTC 324
 Db 277881 TTGTTGACTGGTGAATCTTTTC 277904
 RESULT 15
 AY211269/c 65865 bp DNA linear BCT 21-APR-2003
 LOCUS Campylobacter fetus strain 23D sap gene locus, partial sequence.
 DEFINITION AY211269 AF027405 J05577 L15800 S44580 S78860
 ACCESSION AY211269.1 GI:28974206
 VERSION
 KEYWORDS Campylobacter fetus
 SOURCE Campylobacter fetus
 ORGANISM Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
 Campylobacteraceae; Campylobacter.
 REFERENCE 1 Blaser, M.J. and Gotschlich, E.C.
 TITLE Surface array protein of Campylobacter fetus. Cloning and gene
 structure
 JOURNAL J. Biol. Chem. 265 (24), 14529-14535 (1990)
 MEDLINE 90354448
 PUBMED 2387868
 REFERENCE 2 (bases 1 to 65865)
 AUTHORS Blaser, M.J. and Gotschlich, E.C.
 TITLE Surface array protein of Campylobacter fetus. Cloning and gene

JOURNAL J. Biol. Chem. 265 (31), 19372 (1990)
 MEDLINE 91035477
 PUBMED 2229082
 REFERENCE 3 (bases 1 to 65865)
 AUTHORS Tummuru, M.K. and Blaser, M.J.
 TITLE Characterization of the Campylobacter fetus sapA promoter: evidence
 that the sapA promoter is deleted in spontaneous mutant strains
 JOURNAL J. Bacteriol. 174 (18), 5916-5922 (1992)
 MEDLINE 92394895
 PUBMED 1522068
 REFERENCE 4 (bases 1 to 65865)
 AUTHORS Tummuru, M.K. and Blaser, M.J.
 TITLE Rearrangement of sapA homologs with conserved and variable regions
 in Campylobacter fetus
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 90 (15), 7265-7269 (1993)
 MEDLINE 8346244
 PUBMED 8346244
 REFERENCE 5 (bases 1 to 65865)
 AUTHORS Dworin, J., Tummuru, M.K. and Blaser, M.J.
 TITLE A lipopolysaccharide-binding domain of the Campylobacter fetus
 S-layer protein resides within the conserved N terminus of a family
 of silent and divergent homologs
 JOURNAL J. Bacteriol. 177 (7), 1734-1741 (1995)
 MEDLINE 95204538
 PUBMED 7896695
 REFERENCE 6 (bases 1 to 65865)
 AUTHORS Thompson, S.A., Shedd, O.L., Ray, K.C., Beins, M.H., Jorgensen, J.P. and
 Blaser, M.J.
 TITLE Campylobacter fetus surface layer proteins are transported by a
 type I secretion system
 JOURNAL J. Bacteriol. 180 (24), 6450-6458 (1998)
 MEDLINE 96063317
 PUBMED 9851986
 REFERENCE 7 (bases 1 to 65865)
 AUTHORS Tu, Z.C., Wassenaar, T.M., Thompson, S.A. and Blaser, M.J.
 TITLE Structure and genotypic plasticity of the Campylobacter fetus sap
 locus
 JOURNAL Mol. Microbiol. 48 (3), 685-698 (2003)
 PUBMED 12694614
 REFERENCE 8 (bases 1 to 65865)
 AUTHORS Blaser, M.J.
 TITLE Direct Submission
 JOURNAL Submitted (20-MAY-1993) Division of Infectious Diseases, Vanderbilt
 University School of Medicine, Nashville, TN 37232, USA
 REFERENCE 9 (bases 1 to 65865)
 AUTHORS Thompson, S.A., Shedd, O.L., Ray, K.C., Beins, M.H., Jorgensen, J.P. and
 Blaser, M.J.
 TITLE Direct Submission
 JOURNAL Submitted (30-SEP-1997) Division of Infectious Diseases, Vanderbilt
 University, A-3310 MCN, Nashville, TN 37232, USA
 REFERENCE 10 (bases 1 to 65865)
 AUTHORS Tu, Z.C., Wassenaar, T.M., Thompson, S.A. and Blaser, M.J.
 TITLE Direct Submission
 JOURNAL Submitted (03-JAN-2003) Microbiology and Medicine, New York
 University School of Medicine, VAMC ROOM 6006W 423 E. 23rd Street,
 New York, NY 10016, USA
 COMMENT On or before Mar 15, 2003 this sequence version replaced gi:255285,
 gi:4009445, gi:144211, gi:289533, gi:913763.
 FEATURES
 source Location/Qualifiers
 1..65865
 /organism="Campylobacter fetus"
 /mol_type="genomic DNA"
 /strain="23D"
 /serotype="A"
 /db_xref="taxon:196"
 1..714
 /gene="mtfA"
 <1..714
 /gene="mtfA"
 /note="Cf0001"
 /codon_start=1
 /transl_table=1
 gene
 CDS

```
gene 3585..4769
/ gene="CG10004"
/ gene=".4769"
CDS 3585..4769
/ note="TIGR00611:recf: Recf protein"
/ codon_start=1
/ transl_table=1
/ product="Recombinational DNA repair ATPase"
/ protein_id="BAB97397.1"
/ db_xref="GI:21322768"
/ translation="WHIRLSLELDYRSWPELKYDLEPGITVIGNGREKTIYVAIG
YLAHLSHRVSDAPLFAVRAENARVSAVAVQGRBLAHLIKPNAQAQSLNTRKV
RPRELIGSVKTVLFAVEDLAVKGPASRRRLDIIATROPBAGVADYKVLKQ
RNALTKATATLRGQYTERGAALSTLDTDQGLARLGAEVMAAFALINELGKTY
EAYTTAPESRPAAVYKTTIDQLSQSFSGDAGITELTELAKROREIERGSSL
VGPBRDVIDMDQAPKAPGASGEMTSLIRIAEPNLKSLSDGDPDILIDVSE
LDAGREKLVGIAQVEQVLTAAVHDDPENLKYLTQGHVTVQDTGTGRISLIDV
Qp"
gene 4766..5302
/ gene="CG10005"
/ gene=".5302"
CDS 4766..5302
/ gene="CG10005"
/ codon_start=1
/ transl_table=1
/ product="Hypothetical protein"
/ protein_id="BAB97398.1"
/ db_xref="GI:21322769"
/ translation="MTDPLEQAFERIRAEARRNGSVPLINKDAFRPBPAGVYK
RKGRASGLDRQKRYVGAESLGSVYKKEIGRGGKDIAGMWTSMNEELVGAIA
CHRPVMIKDKKLPICDSYAMATNLMQGRQLQYIAKVGNNITELRIIGQPAS
WRKGPLVYKGRGPRDTYG"
gene 5354..7489
/ gene="CG10006"
/ gene=".7489"
CDS 5354..7489
/ gene="CG10006"
/ EC_number="5.99.1.3"
/ note="PF00204:DNA topoisomerase II (N-terminal region)
PF00986:DNA gyrase B subunit, carboxyl terminus
PF01751:toprim domain
PF02518:Histidine kinase-, DNA gyrase B-, phycochrome-like
ATPase"
/ translation="GREG: DNA topoisomerase IV, B subunit
TIGR01058:pare Gpos: DNA topoisomerase IV, B subunit
TIGR01059:gyrB: DNA gyrase, B subunit"
/ codon_start=1
/ transl_table=1
/ product="DNA gyrase (topoisomerase II) B subunit"
/ protein_id="BAB97399.1"
/ db_xref="GI:21322770"
/ translation="MRGTWGPVKRVMKRLYRISSECSLKVANTEHNYDASITILE
GSAVRPBMVYIGSTGRGLHLIMEVNVSDVAMAGATVEYTLLEDGVCVD
DGRGIPDMHRSQAPVQVYMTOLHAGKGSQSVAGSHGVGSAVNAISTRREA
DIKLGKHWQNFESKYPDELISGNARGTGTTIRPPDAIEETTERPETSRLO
EMALNKGKLTITLNDPATDELELALAQGSTALELSDIDNTELEVEITDAPK
KPKRERKKLFHYENGLEDVYHINRSKTHPSIVSEKGDHVEVAMQNSVYK
ESVATPANTINREGCTHEGFSALISLNRVAREKLEKEANILTGDCCEGSA
VISVADPOPEGQTKTKLGNTEIKSFVQMANEHIGHMEANPAKVIINKAVSA
QARLAARKADILVRKSGATDGLPGCLACRSKDPKSELYIVEGSGAGSKSRD
SMRQAILPLRGKILNVEKARDLVKLAKEAOAITLALGTHIDEPDINKRHYKTYLM
ADADVQGHATLTLTLTFRFMDLVAGHYLAQPLVYTLKWRGPRGAYSDERD
EQNLMEAGAKINKDGIQRYKGLGEMASELMEITMTPTVAILRVDITTDQRAD
LFSLMGDDVVAARSFTTRNAKVRFLDI"
gene 7830..8798
/ gene="CG10007"
/ gene=".8798"
CDS 7830..8798
/ gene="CG10007"
/ codon_start=1
/ transl_table=1
/ product="Hypothetical protein"
/ protein_id="BAB97400.1"
/ db_xref="GI:21322771"
/ translation="MIQTYLIKQSLKMAOREMLKTLITDILGHVSTPTTAHP
LPDGSSTPVQIWMASDNDSQLVMLPFGNGGYRRPLAALNKAQHPVAIGELRGQ
```

```
QSSAKRSKQWGYHDIASVDFPIQIAAKALDLBEGHMRPLSHMSGQISCLPAA
RPEAKYNLRAIRFGAGSPFRFPFSRPMKRLGGLGVLLGGIGCHI VGVMPKVLGK
DLVYGRSGTHMRBMRPHKNSLIDLTDQDINVEYMKVSIPTFSPCPDDECP
QASIDGLASFVPAQIQMTEIPRALGHNRRARBPESVYKLFLEQAL"
gene 8795..9466
/ gene="CG10008"
/ gene=".9466"
CDS 8795..9466
/ gene="CG10008"
/ codon_start=1
/ transl_table=1
/ product="Hypothetical membrane protein"
/ protein_id="BAB97401.1"
/ db_xref="GI:21322772"
/ translation="MNSPNKRPDLPAQRFSLPTRMAFVFGIFIMSPAIGITVAHIG
TTTISLPVYVTAASGLSLGMITVFNFGPMILQIIVLRQPKOMLVQILMAFLGF
LCDLSLQTLTAQTDNFFVAMIVYVTTIMSGIVPQLVNTITFEGISVIALVYK
FPNVEFGMQIVDMTFVSVAALISWITMGSLIGVREGTFAAFIFGFVQRKLYL
RSIGH"
gene 9914..9914
/ gene="CG10009"
/ gene=".9914"
CDS 9914..9914
/ gene="CG10009"
/ codon_start=1
/ transl_table=1
/ product="Hypothetical protein"
/ protein_id="BAB97402.1"
/ db_xref="GI:21322773"
/ translation="MTSLRDLVSTNDAALADNSAFIEFTVGKSGISGIALKANG
AATKIDIDIVAKGSRRLPELIALSDLMQBYONGTAAPFGALLENSSALDAIS
VADNARSINVPGLGKYVGVGRKAATVBOELPALIQLEKNAQ"
gene 110104..11177
/ gene="CG10010"
/ gene=".11177"
CDS 110104..11177
/ gene="CG10010"
/ codon_start=1
/ transl_table=1
/ product="Hypothetical protein"
/ protein_id="BAB97403.1"
/ db_xref="GI:21322774"
/ translation="MFMIAQRTPLPIHTAHPILPVARYFHQIRATDRLSLQRULES"
Query Match 100.0%; Score 324; DB 1; Length 340000;
Best Local Similarity 100.0%; Pred. No. 1.1e-88;
Matches 324; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 ATGACAACTGATTTCTCCTGATTTCTCCTTGTTCGAGATATGAGCATTAATTTT 60
DB 277581 ATGACAACTGATTTCTCCTGATTTCTCCTTGTTCGAGATATGAGCATTAATTTT 277640
61 GCGCTCCGGCGGCTCCGTTCTTAATCCTTAAGCCCTAGCGTAATCAATTTGTGGGC 120
DB 277641 GCGCTCCGGCGGCTCCGTTCTTAATCCTTAAGCCCTAGCGTAATCAATTTGTGGGC 277700
121 AAAAAGCGATGTGATGTCGACAGAGATCCTTGCAATTTGACCGATCAACGTTGGC 180
DB 277701 AAAAAGCGATGTGATGTCGACAGAGATCCTTGCAATTTGACCGATCAACGTTGGC 277760
181 ACCAATGCGATGATCTAAGACTTAACCTTATGCTCATTTGCGGTTGCGATTACAGTG 240
DB 277761 ACCAATGCGATGATCTAAGACTTAACCTTATGCTCATTTGCGGTTGCGATTACAGTG 277820
241 GTGGCGCATCTTCTTGGCGGTCGACGACCTTGTGTGACGGTTGGCGCTGGACCATCTTT 300
DB 277821 GTGGCGCATCTTCTTGGCGGTCGACGACCTTGTGTGACGGTTGGCGCTGGACCATCTTT 277880
301 TTGTGACATGATGATCTTTTTC 324
DB 277881 TTGTGACATGATGATCTTTTTC 277904
RESULT 14
AX120085 349980 bp DNA linear PAT 11-MAY-2001
LOCUS
```

/gene="brnF"
/note="LIV-E transporter component 2; branched-chain amino acid export protein"
/codon_start=1
/transl_table=11
/product="brnF"
/protein_id="AA046686.1"
/db_xref="GI:21311381"
/translation="MKTQELHSLSVSPSKALEPDDKGRYVEIAQELKTSIAGLGMVPIGIAFGVLVIOGYEWMAPLFSGLIIPAGSEMLVIALVGAAPLGAIALTLVNFHVFAYFSPPLVAVKQIPARAFYFALIDEAVATYARAPAGMSAWRLISQIAHSYVWFGTLGVAIAELIPEIKGEPALCSLFTVTLDSCKTKQISILLAGSFTI ALVIRGQALFAALILFIQLTLIRFFIGRAK"
1538..1993
/gene="ltp"
1538..1993
/gene="ltp"
/note="ltp"
/codon_start=1
/transl_table=11
/product="ltp-like regulator"
/protein_id="AA046687.1"
/db_xref="GI:21311382"
/translation="MKLSDIRALIAELSANARISNLADKVLTPGPCLRVQRIEABGIIIGYSADIHPAVMNRGFEVDTLSNPDSTVNSFSSVAQHEVEIHLRLGSPDYFVRIGVADLEAVEQFLSHIQTVPGIAKISRFAMKVVKPARPQV"

ORIGIN

Query Match 100.0%; Score 324; DB 1; Length 2105;
Best Local Similarity 100.0%; Pred. No. 6, 2e-89;
Matches 324; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGACACATGATTTCCCTGATTCCTCTGTTGCGAGATGCGATCACTACTTT 60
DB 711 ATGACACATGATTTCCCTGATTCCTCTGTTGCGAGATGCGATCACTACTTT 652
QY 61 GCGCTCCGGGCGGTCCTGTTTAAATCCTTAAGCCCTTACGTGAATCAATTTGGGCG 120
DB 651 GCGCTCCGGGCGGTCCTGTTTAAATCCTTAAGCCCTTACGTGAATCAATTTGGGCG 592
QY 121 AAAATGGCGATGATGCGACAGAGAAATCCTTGCAATTTGACGCGATCAACGTTTGGC 180
DB 591 AAAATGGCGATGATGCGACAGAGAAATCCTTGCAATTTGACGCGATCAACGTTTGGC 532
QY 181 AGCAATGCGATGATCTGAAGACTTAACCTTGTGCTCAATTCCTGCGATTAACAGTG 240
DB 531 AGCAATGCGATGATCTGAAGACTTAACCTTGTGCTCAATTCCTGCGATTAACAGTG 472
QY 241 GTGGCGCATCTTCTGGCGGCTGAGAGCACTTGTGAGGCGTGGGCGACCATCGTT 300
DB 471 GTGGCGCATCTTCTGGCGGCTGAGAGCACTTGTGAGGCGTGGGCGACCATCGTT 412
QY 301 TTTGTTGACACTGGTGAATCTTTTC 324
DB 411 TTTGTTGACACTGGTGAATCTTTTC 388

RESULT 13
AP005274 340000 bp DNA linear BCT 08-AUG-2002
LOCUS Corynebacterium glutamicum ATCC 13032 DNA, complete genome, section 1/10.
DEFINITION
ACCESSION AP005274 BA000036
VERSION AP005274.1 GI:21322764
KEYWORDS
SOURCE Corynebacterium glutamicum ATCC 13032
ORGANISM Corynebacterium glutamicum ATCC 13032
Bacteria; Actinobacteriales; Actinomycetales;
Corynebacteriaceae; Corynebacterium.
REFERENCE
1 Nakagawa, S.
AUTHORS Complete genomic sequence of Corynebacterium glutamicum ATCC 13032
TITLES
JOURNAL Unpublished

REFERENCE

2 (bases 1 to 340000)
AUTHORS Nakagawa, S.
TITLE Direct Submission
JOURNAL Submitted (24-MAY-2002) Satoshi Nakagawa, Kyowa Hako Kogyo Co. Ltd., Tokyo Research Laboratories; 3-6-6, Asahi-machi, Machida, Tokyo 194-8533, Japan (E-mail: snakagawa@xanagen.com, Tel: 81-44-829-1031, Fax: 81-44-813-1651)

COMMENT

This sequence is conducted by collaboration of Kyowa Hako Kogyo Co. Ltd. And Kitasato University.

FEATURES

Location/Qualifiers
1..340000

/organism="Corynebacterium glutamicum ATCC 13032"
/mol_type="genomic DNA"
/strain="ATCC 13032"
/db_xref="taxon:195627"
/note="ATCC 13032"
1..1575
/gene="Cg10001"
1..1575
/gene="Cg10001"
/note="PF00308: Bacterial dnaA protein
TIGR00362: dnaA: chromosomal replication initiator protein
dnaA"

CDS

gene

/transl_table=1
/product="ATPase involved in DNA replication initiation"
/protein_id="BAB97394.1"
/db_xref="GI:21322765"
/translation="MSQSSSLLETTWROVADLTTLTSGAASGFDLPTORAVMLTKPIALVDGIVLSTPNAKAVTENDLGNALRYVLSLRGSSSLAVSVEPBEIETPAQGERKYOPDAPVSSNKAQYEVGSGEASTSGMERTHAPAPPEHPATADPELATFORIPRETPEHPANPNREVSINPKYFESFVIGPFNFAAAVAASPAKFNPLFISGSGSGKTHLAAVGNVAQELQPGILIKYSSSEFTNDYINSVDDRQETKRRYRNLDLIMVDDIOFLAGKEQOEPEFHFNALQADKOIILSSDRPKQLTLIEDLRFPFGGLITDIOPLDLETRIALIMKAQDGHVREVELIASSESSIRLEBGLIRVSASSLINQPIDEMAIVARLDLIPDEDEITAPVMEVTAEFESISVDTLGA GKTAVAHARQALMTYICRELTDMSLPKIGDFGCKHTTVMYADRLRQEMTKRDTYDEIQQLTOLIKSRGN"

gene

CDS

/transl_table=1
/product="Hypothetical protein"
/protein_id="BAB97395.1"
/db_xref="GI:21322766"
/translation="MIVDFEVLIMIDPKYRILSPVCAVSPAYPOFLGSCAQGKPRGITFPVTKLTQLQVCNTAVAKVVDNFVSAFISYPQNPCTTPKLGKDPEDSSBAAELNPFV"
2292..3476
/gene="Cg10003"
2292..3476
/gene="Cg10003"
/EC_number="2.7.7.7"
/note="PF00712: DNA polymerase III beta subunit, N-terminal domain
PF02767: DNA polymerase III beta subunit, central domain
PF02768: DNA polymerase III beta subunit, C-terminal domain
TIGR00663: danA: DNA polymerase III, beta subunit"

gene

CDS

/transl_table=1
/product="DNA polymerase III beta subunit"
/protein_id="BAB97396.1"
/db_xref="GI:21322767"
/translation="WESQNVSPVAREDDVTAVATVARNIPTEKTPVAVANLITADD EGSALGIDYDVSIVRLSAVSQPKRIAVAGKLSITGSLPNKPVDFRIDSKAEVTCGSGREPLILPLDDVPMPLPKLPAAGGSINALLFEAVSQVSAAGKDSLEMLTGV SMEIVGQIMLAATDRRLALRTFEMENDPEINAVKILIPARTLLDNARSLDGLND IDIAGVGDOVREGGLFGVHTDNRLETTRLLDADFPYIAPLPRKEHTALSVIAPLY DAIIRVSVLERNAGIYIHSBGQVITAGTAGAHEETLPCAFGKELTIANGY LKQGLSVPTSPRAVFGFTSPSRALMIPEDENPANNENGIPTPTDTYTYLLMPVRL PG"

ORIGIN

Query Match 100.0%; Score 324; DB 6; Length 1271;
 Best Local Similarity 100.0%; Pred. No. 5.8e-89;
 Matches 324; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGACACTGATTTTCTCCGATTTCTCTGTTGTCGAGATGAGTGCATTAATTTT 60
 DB 853 ATGACACTGATTTTCTCCGATTTCTCTGTTGTCGAGATGAGTGCATTAATTTT 912
 QY 61 GCGCTCCGGCGGGTCCGTTCTTAATCCCTTAAGCCCTACGTAATCAAAATTTG 120
 DB 913 GCGCTCCGGCGGGTCCGTTCTTAATCCCTTAAGCCCTACGTAATCAAAATTTG 972
 QY 121 AAAATGGCGATGTGATGCCAGAGGAATCCTTGCCATTTTACCCGATCAACGTTTGC 180
 DB 973 AAAATGGCGATGTGATGCCAGAGGAATCCTTGCCATTTTACCCGATCAACGTTTGC 1032
 QY 181 AGCAATGCGATGTGATGCCAGAGGAATCCTTGCCATTTTACCCGATCAACGTTTGC 240
 DB 1033 AGCAATGCGATGTGATGCCAGAGGAATCCTTGCCATTTTACCCGATCAACGTTTGC 1092
 QY 241 GTGGCGCATCTTCTTGCGCGTGCAGCGACCTTGTTGAGCGTTGGCGCTGGCAATCGTT 300
 DB 1093 GTGGCGCATCTTCTTGCGCGTGCAGCGACCTTGTTGAGCGTTGGCGCTGGCAATCGTT 1152
 QY 301 TTTGTTGACTGCTGTAATCTTTTC 324
 DB 1153 TTTGTTGACTGCTGTAATCTTTTC 1176

RESULT 9
 AX137714 1271 bp DNA linear PAT 30-MAY-2001
 LOCUS AX137714
 DEFINITION Sequence 6 from Patent BP1096010.
 ACCESSION AX137714
 VERSION AX137714.1 GI:14273893
 KEYWORDS
 SOURCE Corynebacterium glutamicum
 ORGANISM Corynebacterium glutamicum
 Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 Corynebacteriaceae; Corynebacterium.

REFERENCE
 1 Kernerhnecht, N., Eggeling, L., Sahm, H. and Pfeifferle, W.
 Nucleotide sequences coding for branched-chain amino acids export
 proteins, method for isolating them and their use
 Patent: EP 1096010-A 6 02-MAY-2001;
 Degussa AG (DE); FORSCHUNGSZENTRUM JUELICH GMBH (DE)
 JOURNAL Location/Qualifiers
 FEATURES
 source 1..1271
 /organism="Corynebacterium glutamicum"
 /mol_type="unassigned DNA"
 /db_xref="taxon:1718"
 /note="ATCC13032"
 101..856
 /note="unnamed protein product; brnF"
 /codon_start=1
 /transl_table=1
 /protein_id="CAC39927.1"
 /db_xref="GI:14273894"
 /db_xref="REMBL:CAC39927"
 /translation="MOKTQIHSSLEVSPEKALBPDDKGRREYEIAQGLKTSIAAGL
 GWPRTIGALIVTQVGEWMAAPLEGLIFAGSTEMLVTAIVVGAAPGAILTTLL
 VNRHVTYASFPFHYKNPIARPSVPAIDAVATARPAWGSAMRLISMQIAFH
 SYVFGGTEGVAIMELEPFIKIEPFLCSLPTVLITDSRTKKQPSLLIAGLSFTI
 ALVTPGQALFALLFLGLTIKYPFLGRRAK"
 853..1179
 /note="unnamed protein product; brnE"
 /codon_start=1
 /transl_table=1
 /protein_id="CAC39928.1"
 /db_xref="GI:14273895"
 /db_xref="REMBL:CAC39928"

CDS

ORIGIN

Query Match 100.0%; Score 324; DB 6; Length 1271;
 Best Local Similarity 100.0%; Pred. No. 5.8e-89;
 Matches 324; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGACACTGATTTTCTCCGATTTCTCTGTTGTCGAGATGAGTGCATTAATTTT 60
 DB 853 ATGACACTGATTTTCTCCGATTTCTCTGTTGTCGAGATGAGTGCATTAATTTT 912
 QY 61 GCGCTCCGGCGGGTCCGTTCTTAATCCCTTAAGCCCTACGTAATCAAAATTTG 120
 DB 913 GCGCTCCGGCGGGTCCGTTCTTAATCCCTTAAGCCCTACGTAATCAAAATTTG 972
 QY 121 AAAATGGCGATGTGATGCCAGAGGAATCCTTGCCATTTTACCCGATCAACGTTTGC 180
 DB 973 AAAATGGCGATGTGATGCCAGAGGAATCCTTGCCATTTTACCCGATCAACGTTTGC 1032
 QY 181 AGCAATGCGATGTGATGCCAGAGGAATCCTTGCCATTTTACCCGATCAACGTTTGC 240
 DB 1033 AGCAATGCGATGTGATGCCAGAGGAATCCTTGCCATTTTACCCGATCAACGTTTGC 1092
 QY 241 GTGGCGCATCTTCTTGCGCGTGCAGCGACCTTGTTGAGCGTTGGCGCTGGCAATCGTT 300
 DB 1093 GTGGCGCATCTTCTTGCGCGTGCAGCGACCTTGTTGAGCGTTGGCGCTGGCAATCGTT 1152
 QY 301 TTTGTTGACTGCTGTAATCTTTTC 324
 DB 1153 TTTGTTGACTGCTGTAATCTTTTC 1176

RESULT 10
 BD014991 1271 bp DNA linear PAT 27-AUG-2002
 LOCUS BD014991
 DEFINITION Nucleotide sequence encoding the discharge of branched amino acid,
 method of isolating the same and utilization thereof.
 ACCESSION BD014991
 VERSION BD014991.1 GI:2255798
 KEYWORDS JP 2001169788-A/1.
 SOURCE unidentified
 ORGANISM unidentified

REFERENCE
 1 (bases 1 to 1271)
 Kernerhnecht, N., Sahm, H., Eggeling, L. and Pfeifferle, W.
 Nucleotide sequence encoding the discharge of branched amino acid,
 method of isolating the same and utilization thereof
 Patent: JP 2001169788-A 1 26-JUN-2001;
 DEGUSSA HUELS AG; FORSCHUNGSZENTRUM JUELICH GMBH
 JOURNAL Location/Qualifiers
 COMMENT OS Corynebacterium glutamicum ATCC14752
 PN JP 2001169788-A/1
 PD 26-JUN-2001
 PF 24-OCT-2000 JP 2000324315
 PR 27-OCT-1999 DE 19951708.8
 PI NICOLE KERNERHNECHT, HERMANN SAHM, LOTHAR EGGERING, WALTER PI
 PFEFFERLE
 PC C12N15/09, C07K14/34, C12N1/21, C12P13/06, C12P13/08, C12P13/08, C12N15/09, C12R1/15, C12N1/21, C12R1/15, C12P13/06, C12R1/15, C12P13/06, C12R1/15)
 PC C12N15/00, C12R1/15)
 PC C12N15/00, C12R1/15)
 CC brnF
 CC brnE
 CC key
 CC key
 FT gene Location/Qualifiers
 FT gene (853)..(1176).
 Location/Qualifiers
 1..1271
 /organism="unidentified"
 /mol_type="genomic DNA"
 /db_xref="taxon:32644"

FEATURES
 source 1..1271
 /organism="unidentified"
 /mol_type="genomic DNA"
 /db_xref="taxon:32644"

ORIGIN

RESULT 6
LOCUS AR391953 1271 bp DNA linear PAT 18-DEC-2003
DEFINITION Sequence 1 from patent US 6613545.
ACCESSION AR391953
VERSION AR391953.1 GI:40115724
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE
1 (bases 1 to 1271)
Kenneknecht, N., Sahm, H., Eggeling, L. and Pfeifferle, W.
Nucleotide sequences coding for the export of branched amino
acids, process for the isolation thereof and use thereof
Patent: US 6613545-A 1 02-SEP-2003;
Location/Qualifiers
1..1271
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Query Match 100.0%; Score 324; DB 6; Length 1271;
Best Local Similarity 100.0%; Pred. No. 5.8e-89;
Matches 324; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGACACGATGATTTCCCTGATTCCTGTTGTCGAGTATGACGATCTACTTT 60
DB 853 ATGACACGATGATTTCCCTGATTCCTGTTGTCGAGTATGACGATCTACTTT 912
QY 61 GCGCTCCGGCGGTCCTGTTCTTAATCCTTAAGCCCTACGTAATCACAATTTGTGGC 120
DB 913 GCGCTCCGGCGGTCCTGTTCTTAATCCTTAAGCCCTACGTAATCACAATTTGTGGC 972
QY 121 AAATGCGCATGATGATGCGACAGAGAACTCTTGCCATTTGACCGCATCAAGTTGGC 180
DB 973 AAATGCGCATGATGATGCGACAGAGAACTCTTGCCATTTGACCGCATCAAGTTGGC 1032
QY 181 AGCAATGCGATGATGATGAGACTCTACCTTTGCTCATTTGCCGTTGCAATTCAGTG 240
DB 1033 AGCAATGCGATGATGATGAGACTCTACCTTTGCTCATTTGCCGTTGCAATTCAGTG 1092
QY 241 GTGGCGCATTTTGTGGCGGTGACGACCACTTTGAGGTTGGCGCTGGCAATCGTT 300
DB 1093 GTGGCGCATTTTGTGGCGGTGACGACCACTTTGAGGTTGGCGCTGGCAATCGTT 1152
QY 301 TTGTTGACTGATGATCTTTTC 324
DB 1153 TTGTTGACTGATGATCTTTTC 1176

RESULT 7
LOCUS AR391956 1271 bp DNA linear PAT 18-DEC-2003
DEFINITION Sequence 6 from patent US 6613545.
ACCESSION AR391956
VERSION AR391956.1 GI:40115727
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE
1 (bases 1 to 1271)
Kenneknecht, N., Sahm, H., Eggeling, L. and Pfeifferle, W.
Nucleotide sequences coding for the export of branched chain amino
acids, process for the isolation thereof and use thereof
Patent: US 6613545-A 6 02-SEP-2003;
Location/Qualifiers
1..1271
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Query Match 100.0%; Score 324; DB 6; Length 1271;

Best Local Similarity 100.0%; Pred. No. 5.8e-89;
Matches 324; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGACACGATGATTTCCCTGATTCCTGTTGTCGAGTATGACGATCTACTTT 60
DB 853 ATGACACGATGATTTCCCTGATTCCTGTTGTCGAGTATGACGATCTACTTT 912
QY 61 GCGCTCCGGCGGTCCTGTTCTTAATCCTTAAGCCCTACGTAATCACAATTTGTGGC 120
DB 913 GCGCTCCGGCGGTCCTGTTCTTAATCCTTAAGCCCTACGTAATCACAATTTGTGGC 972
QY 121 AAATGCGCATGATGATGCGACAGAGAACTCTTGCCATTTGACCGCATCAAGTTGGC 180
DB 973 AAATGCGCATGATGATGCGACAGAGAACTCTTGCCATTTGACCGCATCAAGTTGGC 1032
QY 181 AGCAATGCGATGATGATGAGACTCTACCTTTGCTCATTTGCCGTTGCAATTCAGTG 240
DB 1033 AGCAATGCGATGATGATGAGACTCTACCTTTGCTCATTTGCCGTTGCAATTCAGTG 1092
QY 241 GTGGCGCATTTTGTGGCGGTGACGACCACTTTGAGGTTGGCGCTGGCAATCGTT 300
DB 1093 GTGGCGCATTTTGTGGCGGTGACGACCACTTTGAGGTTGGCGCTGGCAATCGTT 1152
QY 301 TTGTTGACTGATGATCTTTTC 324
DB 1153 TTGTTGACTGATGATCTTTTC 1176

RESULT 8
LOCUS AX137709 1271 bp DNA linear PAT 30-MAY-2001
DEFINITION Sequence 1 from Patent EP1096010.
ACCESSION AX137709
VERSION AX137709.1 GI:14273886
KEYWORDS
SOURCE
ORGANISM
Corynebacterium glutamicum
Corynebacterium glutamicum
Bacteria; Actinobacteria; Actinobacteriales; Corynebacteriaceae; Corynebacterium.

REFERENCE
1 Kenneknecht, N., Eggeling, L., Sahm, H. and Pfeifferle, W.
Nucleotide sequences coding for branched-chain amino acids export
Proteins, method for isolating them and their use
Patents: EP 1096010-A 1 02-MAY-2001;
Degussa AG (DE) ; FORSCHUNGSZENTRUM JUELICH GMBH (DE)
Location/Qualifiers
1..1271
/organism="Corynebacterium glutamicum"
/mol_type="unassigned DNA"
/db_xref="taxon:1718"
/note="ATCC14752"
101..856
/note="unnamed protein product; brnF"

CDS

/codon_start=1
/transl_table=1
/protein_id="CAC39923.1"
/db_xref="GI:14273887"
/db_xref="REMBL: CAC39923"
/translation="NOKTOEHSLSLEVSPSKALPPDDKGRYREIAQGLKTSIAGL
GMVPIGALIGLIVIOYGEWMAALFSGLIPAGSTEMLVIALVGAALGIALATTL
VNFHVFARSPFLHVVKNPIARPSVPAIDEDAVYAAAPASMAARLISQIAPH
SYWFGGLTGVAINELFPEIKGEPLCSFYTLITDSCRTKQIPSLLAGSFTI
ALVPIQALPALILFGLITIRYFDGKAK"
853..1179
/note="unnamed protein product; brnE"

CDS

/codon_start=1
/transl_table=1
/protein_id="CAC39924.1"
/db_xref="GI:14273888"
/db_xref="REMBL: CAC39924"
/translation="MTIDFSCILIVAAVITPALRAVPLIILKPIRESQFVKMA
WMPAGILAILTASFRSNAIDKTLIFGLINAVATVAHLIGSRITLISVAGITVFV
GLVNLF"

QY 301 TTGTGGAGCTGGTGAATCTTTTC 324
 DB 301 TTGTGGAGCTGGTGAATCTTTTC 324

RESULT 4
 LOCUS BD014993
 DEFINITION BD014993 324 bp DNA linear PAT 27-AUG-2002
 Nucleotide sequence encoding the discharge of branched amino acid,
 method of isolating the same and utilization thereof.
 ACCESSION BD014993.1 GI:22555800
 VERSION BD014993.1 GI:22555800
 KEYWORDS JP 2001169788-A/3.
 SOURCE unidentified
 ORGANISM unidentified
 unclassified.

REFERENCE
 AUTHORS Kemerukunehito, N., Sahn, H., Eggering, L. and Pfeifferle, W.
 TITLE Nucleotide sequence encoding the discharge of branched amino acid,
 method of isolating the same and utilization thereof
 JOURNAL Patent: JP 2001169788-A 3 26-JUN-2001;
 DEGUSA HUELS AG, FORSCHUNGSZENTRUM JOEHLICH GMBH
 COMMENT OS Corynebacterium glutamicum ATCC14752
 PN JP 2001169788-A/3
 PD 26-JUN-2001
 PF 24-OCT-2000 JP 2000324315
 PR 24-OCT-1999 DE 19851708.8
 PI NICOLE KEMERUKUNEHITO, HERMANN SAHM, LOTHAR EGGERING, WALTER PI
 PFEIFFERLE
 PC C12N15/09, C07K14/34, C12N1/21, C12P13/06, C12P13/08, C12P13/08//
 PC (C12N15/09, C12R1:15), (C12N1/21, C12R1:15), (C12P13/06, C12R1:15),
 PC C12N15/00,
 PC (C12N15/00, C12R1:15)
 CC bme
 CC key
 FT CDS Location/Qualifiers
 1..324 Location/Qualifiers
 1..324 /organism="unidentified"
 /mol_type="genomic DNA"
 /db_xref="taxon:32644"

ORIGIN
 Query Match 100.0%; Score 324; DB 6; Length 324;
 Best Local Similarity 100.0%; Pred. No. 5e-89;
 Matches 324; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGACAACTGATTTCTCCGATATTCCTGTTGCGAGTATGTCAGTCACTTCTTT 60
 DB 1 ATGACAACTGATTTCTCCGATATTCCTGTTGCGAGTATGTCAGTCACTTCTTT 60
 QY 61 GCGCTCCGGGCGGTTCCGTTCTTAATCCTTAAGCCCCCTAAGTGAATCAAAATTGTGGC 120
 DB 61 GCGCTCCGGGCGGTTCCGTTCTTAATCCTTAAGCCCCCTAAGTGAATCAAAATTGTGGC 120
 QY 121 AAATGGCGATGTGAGTGCAGAGCAAGAACTCTTGCAATTTGACCGGATTAAGCTTCCG 180
 DB 121 AAATGGCGATGTGAGTGCAGAGCAAGAACTCTTGCAATTTGACCGGATTAAGCTTCCG 180
 QY 121 AAATGGCGATGTGAGTGCAGAGCAAGAACTCTTGCAATTTGACCGGATTAAGCTTCCG 180
 DB 121 AAATGGCGATGTGAGTGCAGAGCAAGAACTCTTGCAATTTGACCGGATTAAGCTTCCG 180
 QY 181 AGCAATGCGATGATCTGAAGACTCTAACCCTTGGTCTCAATGCCGTTGGATTAACATG 240
 DB 181 AGCAATGCGATGATCTGAAGACTCTAACCCTTGGTCTCAATGCCGTTGGATTAACATG 240
 QY 241 GTGGCGCATCTTTCTGGCGGTGCAGACCACTTGTGAGCGTTGGCGTGCACCAATCGTT 300
 DB 241 GTGGCGCATCTTTCTGGCGGTGCAGACCACTTGTGAGCGTTGGCGTGCACCAATCGTT 300
 QY 301 TTGTGGAGCTGGTGAATCTTTTC 324
 DB 301 TTGTGGAGCTGGTGAATCTTTTC 324

RESULT 5
 LOCUS BD162491
 DEFINITION BD162491 324 bp DNA linear PAT 17-JAN-2003
 Novel polynucleotide.
 ACCESSION BD162491.1 GI:27868249
 VERSION BD162491.1 GI:27868249
 KEYWORDS JP 2002191370-A/290.
 SOURCE unidentified
 ORGANISM unidentified
 unclassified.

REFERENCE
 AUTHORS Nakagawa, S., Mizoguchi, H., Ando, S., Hayashi, M., Ochiai, K.,
 Yokoi, H., Tateishi, N., Senoo, A., Ikeda, M. and Ozaki, A.
 TITLE Novel polynucleotide
 JOURNAL Patent: JP 2002191370-A 290 09-JUL-2002;
 KYOMA HAKKO KOGYO CO LTD
 COMMENT OS Corynebacterium glutamicum
 PN JP 2002191370-A/290
 PD 09-JUL-2002
 PF 15-DEC-2000 JP 2000405096
 PI SATOSHI NAKAGAWA, HIROSHI MIZOGUCHI, SEIKO ANDO, MIKIO HAYASHI,
 PI KEIKO OCHIAI,
 PI HARUHIKO YOKOI, NAKO TATEISHI, AKIHIRO SENOO, MASATO IKEDA, AKIO
 PI OZAKI
 PC C12N15/09, C12N15/09, C07K14/34, C07K16/12, C07K16/40, C12N1/00, PC
 C12N1/15,
 PC C12N1/19, C12N1/21, C12N5/10, C12N9/00, C12N9/02, C12P7/40, C12P13/
 PC 04, C12P13/08,
 PC C12P13/00, C12P13/34, C12P21/02, C12Q1/37, C12Q1/68, G01N33/53, PC
 G01N33/56,
 PC G01N33/56, G01N33/68, G01N37/00, C12P21/08, C12N1/21, C12R1:15,
 PC (C12N1/21, C12R1:13), (C12N1/21, C12R1:01), (C12P13/08, C12R1:15),
 PC C12N15/00,
 PC C12N5/00, C12N15/00
 CC Novel polynucleotide
 CC key
 FT CDS Location/Qualifiers
 1..324 Location/Qualifiers
 1..324 /organism="Corynebacterium glutamicum".
 1..324 /organism="unidentified"
 /mol_type="genomic DNA"
 /db_xref="taxon:32644"

ORIGIN
 Query Match 100.0%; Score 324; DB 6; Length 324;
 Best Local Similarity 100.0%; Pred. No. 5e-89;
 Matches 324; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGACAACTGATTTCTCCGATATTCCTGTTGCGAGTATGTCAGTCACTTCTTT 60
 DB 1 ATGACAACTGATTTCTCCGATATTCCTGTTGCGAGTATGTCAGTCACTTCTTT 60
 QY 61 GCGCTCCGGGCGGTTCCGTTCTTAATCCTTAAGCCCCCTAAGTGAATCAAAATTGTGGC 120
 DB 61 GCGCTCCGGGCGGTTCCGTTCTTAATCCTTAAGCCCCCTAAGTGAATCAAAATTGTGGC 120
 QY 121 AAATGGCGATGTGAGTGCAGAGCAAGAACTCTTGCAATTTGACCGGATTAAGCTTCCG 180
 DB 121 AAATGGCGATGTGAGTGCAGAGCAAGAACTCTTGCAATTTGACCGGATTAAGCTTCCG 180
 QY 121 AAATGGCGATGTGAGTGCAGAGCAAGAACTCTTGCAATTTGACCGGATTAAGCTTCCG 180
 DB 121 AAATGGCGATGTGAGTGCAGAGCAAGAACTCTTGCAATTTGACCGGATTAAGCTTCCG 180
 QY 181 AGCAATGCGATGATCTGAAGACTCTAACCCTTGGTCTCAATGCCGTTGGATTAACATG 240
 DB 181 AGCAATGCGATGATCTGAAGACTCTAACCCTTGGTCTCAATGCCGTTGGATTAACATG 240
 QY 241 GTGGCGCATCTTTCTGGCGGTGCAGACCACTTGTGAGCGTTGGCGTGCACCAATCGTT 300
 DB 241 GTGGCGCATCTTTCTGGCGGTGCAGACCACTTGTGAGCGTTGGCGTGCACCAATCGTT 300
 QY 301 TTGTGGAGCTGGTGAATCTTTTC 324
 DB 301 TTGTGGAGCTGGTGAATCTTTTC 324

FEATURES
 source
 Location/Qualifiers
 1..324
 /organism="unknown"
 /mol_type="genomic DNA"

ORIGIN
 Query Match 100.0%; Score 324; DB 6; Length 324;
 Best Local Similarity 100.0%; Pred. No. 5e-89;
 Matches 324; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 ATGACAACTGATTTCTCCTGATATTCCTGTTGTGCGAGATATGAGATCACTTACTTTT 60
 1 ATGACAACTGATTTCTCCTGATATTCCTGTTGTGCGAGATATGAGATCACTTACTTTT 60
 61 GCGCTCCGGCGGCTCCGTTCTTAATCCCTTAAGCCCTACGTAATCAATTTGTGGGC 120
 61 GCGCTCCGGCGGCTCCGTTCTTAATCCCTTAAGCCCTACGTAATCAATTTGTGGGC 120
 121 AAAATGGCGATGTGATGACGAGAGATCTTGCAATTTGACCGCATCAAGCTTTGCG 180
 121 AAAATGGCGATGTGATGACGAGAGATCTTGCAATTTGACCGCATCAAGCTTTGCG 180
 181 AGCAATGCGATGATCTGAAAGACTTACCTTTGTTCTCATTTGCCGTTACAGTG 240
 181 AGCAATGCGATGATCTGAAAGACTTACCTTTGTTCTCATTTGCCGTTACAGTG 240
 241 GTGGCGCATCTTTGCGCGGTGACGACGACCTTTGTGAGCGTTGGCGCTGACCATCTT 300
 241 GTGGCGCATCTTTGCGCGGTGACGACGACCTTTGTGAGCGTTGGCGCTGACCATCTT 300
 301 TTGTGACTGTGATCTTTTC 324
 301 TTGTGACTGTGATCTTTTC 324

RESULT 2
 AX120374 324 bp DNA linear PAT 11-MAY-2001
 LOCUS
 DEFINITION
 Sequence 290 from Patent EP1108790.
 AX120374
 VERSION
 AX120374.1 GI:14037089
 KEYWORDS
 Corynebacterium glutamicum
 Corynebacterium glutamicum
 Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 Corynebacterineae; Corynebacteriaceae; Corynebacterium.

REFERENCE
 AUTHORS
 Nakagawa, S., Mizoguchi, H., Ando, S., Hayashi, M., Ochiai, K.,
 Yokoi, H., Tateishi, N., Senoh, A., Ikeda, M. and Ozaki, A.
 TITLE
 Novel polynucleotides
 Patents: EP 1108790-A 290 20-JUN-2001;
 KYOWA HAKKO KOSYO CO., LTD. (JP)
 JOURNAL
 Location/Qualifiers
 1..324
 /organism="Corynebacterium glutamicum"
 /mol_type="unassigned DNA"
 /db_xref="taxon:1718"

ORIGIN
 Query Match 100.0%; Score 324; DB 6; Length 324;
 Best Local Similarity 100.0%; Pred. No. 5e-89;
 Matches 324; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 ATGACAACTGATTTCTCCTGATATTCCTGTTGTGCGAGATATGAGATCACTTACTTTT 60
 1 ATGACAACTGATTTCTCCTGATATTCCTGTTGTGCGAGATATGAGATCACTTACTTTT 60
 61 GCGCTCCGGCGGCTCCGTTCTTAATCCCTTAAGCCCTACGTAATCAATTTGTGGGC 120
 61 GCGCTCCGGCGGCTCCGTTCTTAATCCCTTAAGCCCTACGTAATCAATTTGTGGGC 120
 121 AAAATGGCGATGTGATGACGAGAGATCTTGCAATTTGACCGCATCAAGCTTTGCG 180
 121 AAAATGGCGATGTGATGACGAGAGATCTTGCAATTTGACCGCATCAAGCTTTGCG 180

Db 121 AAAATGGCGATGTGATGACGAGAGATCTTGCAATTTGACCGCATCAAGCTTTGCG 180
 QY 181 AGCAATGCGATGATCTGAAAGACTTACCTTTGTTCTCATTTGCCGTTACAGTG 240
 Db 181 AGCAATGCGATGATCTGAAAGACTTACCTTTGTTCTCATTTGCCGTTACAGTG 240
 QY 241 GTGGCGCATCTTTGCGCGGTGACGACGACCTTTGTGAGCGTTGGCGCTGACCATCTT 300
 Db 241 GTGGCGCATCTTTGCGCGGTGACGACGACCTTTGTGAGCGTTGGCGCTGACCATCTT 300
 QY 301 TTGTGACTGTGATCTTTTC 324
 Db 301 TTGTGACTGTGATCTTTTC 324

RESULT 3
 AX137712 324 bp DNA linear PAT 30-MAY-2001
 LOCUS
 DEFINITION
 Sequence 4 from Patent EP1096010.
 AX137712
 VERSION
 AX137712.1 GI:14273891
 KEYWORDS
 Corynebacterium glutamicum
 Corynebacterium glutamicum
 Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 Corynebacterineae; Corynebacteriaceae; Corynebacterium.

REFERENCE
 AUTHORS
 Kernerkecht, N., Eggeling, J., Sahm, H. and Pfeifferle, W.
 TITLE
 Nucleotide sequences coding for branched-chain amino acids export
 proteins, method for isolating them and their use
 Patents: EP 1096010-A 4 02-MAY-2001;
 Degussa AG (DE); FORSCHUNGSZENTRUM JUELICH GMBH (DE)
 JOURNAL
 Location/Qualifiers
 1..324
 /organism="Corynebacterium glutamicum"
 /mol_type="unassigned DNA"
 /db_xref="taxon:1718"
 /note="ATCC14752"
 1..324
 /note="unnamed protein product; brnE"
 /codon_start=1
 /transl_table=1
 /protein_id="CAC39926.1"
 /db_xref="GI:14273892"
 /db_xref="EMBL:CA39926"
 /translation="MTDPFSCILIVAVCAITPFAALRAVPLILKPLRESQFVGKAM
 WMPGILALITASTFRSNAIDKLTITGLAVALTIVVAHLGGRRLISVAGATIVV
 GLVNLV"

ORIGIN
 Query Match 100.0%; Score 324; DB 6; Length 324;
 Best Local Similarity 100.0%; Pred. No. 5e-89;
 Matches 324; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 ATGACAACTGATTTCTCCTGATATTCCTGTTGTGCGAGATATGAGATCACTTACTTTT 60
 1 ATGACAACTGATTTCTCCTGATATTCCTGTTGTGCGAGATATGAGATCACTTACTTTT 60
 61 GCGCTCCGGCGGCTCCGTTCTTAATCCCTTAAGCCCTACGTAATCAATTTGTGGGC 120
 61 GCGCTCCGGCGGCTCCGTTCTTAATCCCTTAAGCCCTACGTAATCAATTTGTGGGC 120
 121 AAAATGGCGATGTGATGACGAGAGATCTTGCAATTTGACCGCATCAAGCTTTGCG 180
 121 AAAATGGCGATGTGATGACGAGAGATCTTGCAATTTGACCGCATCAAGCTTTGCG 180
 181 AGCAATGCGATGATCTGAAAGACTTACCTTTGTTCTCATTTGCCGTTACAGTG 240
 181 AGCAATGCGATGATCTGAAAGACTTACCTTTGTTCTCATTTGCCGTTACAGTG 240
 241 GTGGCGCATCTTTGCGCGGTGACGACGACCTTTGTGAGCGTTGGCGCTGACCATCTT 300
 241 GTGGCGCATCTTTGCGCGGTGACGACGACCTTTGTGAGCGTTGGCGCTGACCATCTT 300

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 19, 2004, 10:57:18 ; Search time 1373.24 Seconds
(without alignments)
10226.291 Million cell updates/sec

Title: US-10-608-504-4

Perfect score: 324
1 atgacaactgattctctctg.....ttgacacggtgacatctttc 324

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 2167151695 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

```

1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_ov:*
5: gb_ov:*
6: gb_ov:*
7: gb_ov:*
8: gb_ov:*
9: gb_ov:*
10: gb_ov:*
11: gb_ov:*
12: gb_ov:*
13: gb_ov:*
14: gb_ov:*
15: gb_ov:*
16: gb_ov:*
17: gb_ov:*
18: gb_ov:*
19: gb_ov:*
20: gb_ov:*
21: gb_ov:*
22: gb_ov:*
23: gb_ov:*
24: gb_ov:*
25: gb_ov:*
26: gb_ov:*
27: gb_ov:*
28: gb_ov:*
29: gb_ov:*
30: gb_ov:*
31: gb_ov:*
32: gb_ov:*
33: gb_ov:*
34: gb_ov:*
35: gb_ov:*
36: gb_ov:*
37: gb_ov:*
38: gb_ov:*
39: gb_ov:*
40: gb_ov:*
41: gb_ov:*

```

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	324	100.0	324	6 AR391955	AR391955 Sequence
2	324	100.0	324	6 AX120374	AX120374 Sequence
3	324	100.0	324	6 AX137712	AX137712 Sequence
4	324	100.0	324	6 BD014993	BD014993 Nucleotide
5	324	100.0	324	6 BD162491	BD162491 Novel pol
6	324	100.0	1271	6 AR391953	AR391953 Sequence
7	324	100.0	1271	6 AR391956	AR391956 Sequence
8	324	100.0	1271	6 AX137709	AX137709 Sequence
9	324	100.0	1271	6 AX137714	AX137714 Sequence
10	324	100.0	1271	6 BD014991	BD014991 Nucleotide
11	324	100.0	1271	6 BD014994	BD014994 Nucleotide
12	324	100.0	2105	1 AF454053	AF454053 Corynebact
13	324	100.0	340000	1 AP005274	AP005274 Corynebact
14	324	100.0	349980	6 AX120085	AX120085 Sequence
15	38.2	11.8	65865	1 AY211269	AY211269 Campyloba
16	36.6	11.3	4534	1 CFU25133	CFU25133 Campylobact
17	36.6	11.1	141605	2 AC013732	AC013732 Homo sapi
18	36.6	11.1	148332	9 AC079586	AC079586 Homo sapi
19	35.8	11.0	2040	8 NT289862	NT289862 Nicotiana
20	35.6	11.0	137678	8 AP005183	AP005183 Oryza sat
21	35.6	11.0	173270	8 AC024886	AC024886 Homo sapi
22	35.6	11.0	328197	2 AC117392	AC117392 Homo sapi
23	35.4	10.9	113025	2 AC109596	AC109596 Oryza sat
24	35.2	10.9	34494	3 AC006633	AC006633 Caenorhab
25	34.8	10.7	2000	6 AX655393	AX655393 Sequence
26	34.8	10.7	86524	4 AC104481	AC104481 Sus scrofa
27	34.8	10.7	270748	2 AC121736	AC121736 Rattus no
28	34.6	10.7	110000	2 AC095209_1	Continuation (2 of
29	34.6	10.7	170475	2 AC136263	AC136263 Rattus no
30	34.6	10.7	179141	2 AC108890	AC108890 Bos tauru
31	34.6	10.7	224691	2 AC137364	AC137364 Rattus no
32	34.6	10.7	270850	2 AC095807	AC095807 Homo sapi
33	34.4	10.6	189805	9 AC004169	AC004169 Homo sapi
34	34.4	10.6	190000	2 AC004479	AC004479 Homo sapi
35	34.2	10.6	1686	1 NC000001	NC000001 Escherichia
36	34.2	10.6	12883	1 AE009221	AE009221 Agrobacte
37	34.2	10.6	12909	1 AE008188	AE008188 Agrobacte
38	34.2	10.6	13243	1 AE004602	AE004602 Pseudomon
39	34.2	10.5	150662	2 AC138550	AC138550 Danto rer
40	34	10.5	167753	2 AC108752	AC108752 Homo sapi
41	34	10.5	168332	2 AC053483	AC053483 Homo sapi
42	34	10.5	169030	2 AC125625	AC125625 Homo sapi
43	34	10.5	253370	5 AL929533	AL929533 Zebrafish
44	33.8	10.4	117645	8 AC144644	AC144644 Medicago
45	33.8	10.4	210791	9 CNS01D17	AL133467 Human chr

ALIGNMENTS

RESULT 1
AR391955
LOCUS AR391955 324 bp DNA linear PAT 18-DEC-2003
DEFINITION Sequence 4 from patent US 6613545.
ACCESSION AR391955
VERSION AR391955.1 GI:40115726
KEYWORDS
SOURCE Unknown.
ORGANISM Unclasseified.

REFERENCE 1 (bases 1 to 324)
Kernertnecht, N., Salm, H., Eggeling, L. and Pfeifferle, W.
Nucleotide sequences coding for the export of branched chain amino
acids, process for the isolation thereof and use thereof
Patent: US 6613545-A 4 02-SEP-2003;
JOURNAL

Db 66 GGANLASSILLTLLTSSQILVGLSMRPLRSQPLRMRLGLFLLTDEFFALTSSQYDRQ 125
QY 147 MSANRLISMQIAFHSYVWGGITGVIAIAELIP-FRIKLEFALCSLFYTLTLDSCRTKKQ 205
Db 126 FNRWYALGVLPFYIAMNFTLIGLIGLPHLDQFGLDPSIVATFVA----- 174
QY 206 IPSLLAGLSFTIALVVIPOALFAAL-----LIFGLTIRYFPL 246
Db 175 ----LIAPLVRLATVVCVAVSLFCSVLSFVWMEFALVAVAGLIGMSAGFI 221

RESULT 15

QYVI98 PRELIMINARY; PRT; 231 AA.
AC QYVI98;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN HH0710.
OS Helicobacter hepaticus.
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Helicobacteraceae; Helicobacter.
OX NCBI_TaxID=32025;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 51449 / 3B1;
RX MEDLINE=22709201; PubMed=12810954;
RA Suerbaum S., Josephans C., Sterzenbach T., Drescher B., Brandt P.,
RA Bell M., Droege M., Fartmann B., Fischer H.-P., Ge Z., Hoerster A.,
RA Holland K., Klein K., Koenig J., Macko L., Mendz G.L., Nyakatura G.,
RA Schauer D.B., Shen Z., Weber J., Wroch W., Fox J.G.,
RT "The complete genome sequence of the carcinogenic bacterium
RT Helicobacter hepaticus.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:7901-7906(2003).
DR EMBL: AB017146; AAP7307.1; -;
KW Hypothetical protein, Complete proteome.
SQ SEQUENCE 231 AA; 25674 MW; 10CDA631C71EP9BD CRC64;

Query Match 15.3%; Score 194; DB 16; Length 231;
Best Local Similarity 26.3%; Pred. No. 7.2e-08;
Matches 51; Conservative 44; Mismatches 85; Indels 14; Gaps 6;

QY 44 LGWPIGIAFGLVIOGY-EMWAAPLFGSLFAGSTEMVIALVGAAPLGAIALTLTL 102
Db 23 LGYIFWGAVFGLLOKTKYGAIMWA-IMAVAIYGGITQFAVGLIASGAGLMSFVLVSM 81
QY 103 VNFHVFYAFSPFLHVVKNPIARFYSVFALIDAVAVTAAR-----PAGWSAMRLISMQ 156
Db 82 INARQIFYSISMERFMRGKKRYMYISLDELALINLKSPRTGVKQWPFDFIAALN 141
QY 157 IAFHSYVWGGITGVIAIAELIPFRIKLEFALCSLFYTLTLDSCRTK-KQIPSLLAGLS 215
Db 142 ---QSYVVIQCTTGAAALGGKLEFNSGLDFWMSAIFVITFEGMRNKTMRIPALI--GIT 196
QY 216 FTIALVVIPOALF 229
Db 197 ISLICLYIFGAGQF 210

Search completed: April 16, 2004, 06:55:50
Job time : 53.6407 secs

06 Escherichia coli O6.
07 Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
08 Enterobacteriaceae; Escherichia.
09 NCBI_TaxID=217992;
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833

DR InterPro:IPR00524; HTH_GntR.
DR Pfam: PF03591; AzC; 1.
DR PROSITE: PS00043; HTH_GNTR_FAMILY; 1.
SQ NON TER 1 1
SQ SEQUENCE 247 AA; 26103 MW; 8958101295841873 CRC64;
Query Match 15.5%; Score 197; DB 2; Length 247;
Best Local Similarity 28.4%; Pred. No. 4.4e-08;
Matches 67; Conservative 42; Mismatches 107; Indels 20; Gaps 7
Qy 13 VSPSKALBPDDKGRRYEINQGLKTSIAQLGMYPIGLIFGLIVIGYEMNAAPLPSG 72
Db 3 VSSSQ---DKPTADASFSAGLPDSLPYIGYVPAFAFGNAVVLGTPDEATLSC 58
Qy 73 LIPAGSEMLVIALVGAAPLGAIALTTLLVFRHVFAFGFPLHVKN-PIAR-FYSVF 130
Db 59 ILYAGASQVITALLSAGASIVVAALITVMADVHVHLYGSLRRRIQRLPFGKTMWAP 118
Qy 131 ALIDEAVATAARPA---GMSAMRLISMOIAFHSYVWFGSLTVAVL---AELIPEIK 182
Db 119 GLIDVFAAATARLARLNDNRNRPMLNGLVLSAMLSWAGVITAVVRNGLPGYVEA 178
Qy 183 GLFPA LCSIFVTLTLDSCRTKKQ-IPSLLLAGLSFTIALVVI PQALFAALLIFLG 237
Db 179 ALAFMPLALFLSFLSLASFRRQSLVVAALGACIGLVSEIF---AAILIGIG 229
RESULT 14
Q88KAS PRELIMINARY; PRT; 230 AA.
AC Q88KAS;
DT 01-JUN-2003 (Tremblrel. 24, Created)
DT 01-JUN-2003 (Tremblrel. 24, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE Branched-chain amino acid transport protein AzC.
DE AzC OR PP2385.
OS Pseudomonas putida (strain KT2440).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID:160488;
RX [1]
RN RP
RS SEQUENCE FROM N.A.
RX MEDLINE:22423060; PubMed:12534463;
RA Nelson K.E., Meinel C., Paulsen I.T., Dodson R.J., Hilbert H.,
RA Martins dos Santos V.A.P., Fouts D.E., Gill S.R., Pop M., Holmes M.,
RA Brinkac L., Benson M., DeBoy R.T., Daugherty S., Kolonay J.,
RA Madupu R., Nelson W., White O., Peterson J., Khouri H., Hance I.,
RA Chitua Lee P., Holtapple E., Scanlan D., Tran K., Moazzaz A.,
RA Urethack T., Rizzo M., Lee K., Kosack D., Moestl D., Wiedler H.,
RA Lauber U., Stjepandic D., Hohnselt J., Streitz M., Heim S.,
RA Kiewitz C., Eisen J., Timmis K.N., Duesterhoeft A., Tuenmler B.,
RA Fraser C.M.;
RT "Complete genome sequence and comparative analysis of the
RT metabolically versatile Pseudomonas putida KT2440.";
RL Environ. Microbiol. 4:799-808(2002).
DR EMBL; AB016783; AAN67998.1; -.
DR TIGR; PP2385; -.
DR GO; GO:0016021; C:Integral to membrane; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro: IPR004471; AzC.
DR Pfam: PF03591; AzC; 1.
KW Complete proteome.
SO SEQUENCE 230 AA; 24752 MW; 320878FDCD403A81 CRC64;
Query Match 15.3%; Score 194; DB 16; Length 230;
Best Local Similarity 25.5%; Pred. No. 7.1e-08;
Matches 59; Conservative 43; Mismatches 101; Indels 28; Gaps 5
Qy 29 RYSIAQGLKTSIAAGLGMYPGIGLIVIGYEMNAAPLPSGLIFAGSEMLVITLV 88
Db 6 RQAFHGAIALILPLSLAVAFNGLLAGSAIEANISAMQGLSIVVAGAAQVALISM LK 65
Qy 89 GAAPLGAIALITLLVFRHVFAFGFPLHVKNPIA-RFYSVFALIDEAVATAAR-AG 146

QY 150 WRISMOIAHFWYVFGGLTGVAILIP-PEIKGIEFALCSFTLTLDSCRTKKQIPS 208
 DB 136 WYALGAGLSFYLTWNLTATLTGIVAGSLIPELNEMGLEFAVAAFTAVPTIKNAVLSS 195
 QY 209 ILLAGISFTIALV--IPGALFALLIFL 236
 DB 196 LVVA-LVGSVALTYVYEGSGLMSITLAWL 224

RESULT 10

Q8DV59 PRELIMINARY; PRT; 232 AA.
 AC Q8DV59;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Putative integral membrane protein, possible branched-chain amino acid
 DE permease.
 DE SMU.388.
 OS Streptococcus mutans.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1309;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=UAI59 / ATCC 700610 / Serotype C;
 RX MEDLINE=2225063; PubMed=12397186;
 RA Adic D., Moshan W.M., McClughlin R.E., Savic G., Chang J.,
 RA Carson M.B., Primeaux C., Tian R., Kenton S., Jia H., Lin S., Qian Y.,
 RA Li S., Zhu H., Najaf F., Lai H., White J., Roe B.A., Peretti J.J.;
 RT "Genome sequence of Streptococcus mutans UAI59, a cariogenic dental
 RT pathogen."
 RL Proc. Natl. Acad. Sci. U.S.A. 99:14434-14439(2002).
 DR EMBL; AEO14886; AANS8144.1; -
 DR GO; GO:0016021; C:Integral to membrane; IEA.
 DR GO; GO:0006810; P:Transport; IEA.
 DR InterPro; IPRO004471; AzIC.
 DR InterPro; IPRO00437; Prok_Lipoprot_S.
 DR Pfam; PF03591; AzIC; 1.
 DR ProSite; PS00013; PROKAR_LIPOPROTEIN; 1.
 KM Complete proteome.
 SQ SEQUENCE 232 AA; 24937 MW; 54654370CC0B47F6 CRC64;

Query Match 15.8%; Score 200.5; DB 16; Length 232;
 Best Local Similarity 28.7%; Pred. No. 2.2e-08;
 Matches 70; Conservative 44; Mismatches 91; Indels 39; Gaps 10;

QY 24 DKGYRREINQKTSIAAGLWYPIGIAFGLLVIOGYEWMAAPLFSGLIRAGSTEMLV 83
 DB 3 EKGEK---EGVAKSLPTLGYVSGIAGGVGANSGLTPFOMGLMSLTVAGSAQFWM 57
 QY 84 IALVGAAPGATLTLTVNFRHVFVAFSPPLAV---VKNPVAFVSVPALI-DEAYA 138
 DB 58 CAMFVAGADPSIYMTVFILNLRN---FLMSLHATITFTKSLWQICIGTLIDESYG 113
 QY 139 VTAAR-----PAGSAMRLISMOIAHFWYVFGGLTGVAILIPF-EIIGLEFALCS 190
 DB 114 VLNEHYHAKOISTAMWAGNNITGYLA---WLFNAILGTALGSVLPENETIGLDPALIA 169
 QY 191 LFTVL-----TLDSCKTKQIPSLLAGISFTIALVVIPOGALFALLIFLG-----L 238
 DB 170 MFWISFESQLAAMQVFKKIGILLAVTLSTYFL-LVIVISLSLAVLSTIGCFAYL 228

QY 239 LTR 242
 DB 229 LDVA 232

RESULT 11
 Q8X308 PRELIMINARY; PRT; 245 AA.
 ID Q8X308;
 AC Q8X308;

DT 01-MAR-2002 (TrEMBLrel. 20, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Oit, hypothetical protein.
 DE 23983 OR EC3544.
 OS Escherichia coli O157:H7.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.
 OX NCBI_TaxID=83334;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=O157:H7 / EDL933 / ATCC 700927;
 RX MEDLINE=21074935; PubMed=11206551;
 RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
 RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
 RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
 RA Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Poranousis K.,
 RA Apodaca U., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
 RA Welch R.A., Blattner F.R.;
 RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7."
 RL Nature 409:529-533(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=O157:H7 / RIMD 0509952;
 RX MEDLINE=21156231; PubMed=11258796;
 RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
 RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tode T.,
 RA Iida T., Takami H., Honda T., Sasaki K., Ogasawara N., Yasunaga T.,
 RA Kikuta S., Shiba T., Hattori M., Shitagawa H.;
 RT "Complete genome sequence of enterohaemorrhagic Escherichia coli
 RT O157:H7 and genomic comparison with a laboratory strain K-12."
 RL DNA Res. 8:111-122(2001).
 DR EMBL; AEO05497; AAG57791.1; -
 DR EMBL; AP002562; BMB36967.1; -
 DR PIR; C89916; C89916.
 DR PIR; H91071; H91071.
 DR GO; GO:0016021; C:Integral to membrane; IEA.
 DR GO; GO:0006810; P:Transport; IEA.
 DR InterPro; IPRO004471; AzIC.
 DR Pfam; PF03591; AzIC; 1.
 KM Complete proteome.
 SQ SEQUENCE 245 AA; 26093 MW; 20AAFD291E644DB3 CRC64;

Query Match 15.6%; Score 198; DB 16; Length 245;
 Best Local Similarity 29.1%; Pred. No. 3.6e-08;
 Matches 68; Conservative 41; Mismatches 95; Indels 30; Gaps 9;

QY 34 QGLKTSIAAGLWYPIGIAFGLLVIOGYEWMAAPLFSGLIRAGSTEMLVIALVGAAPL 93
 DB 17 ECKDLSLPYITISTIPAFAGLNATLGRSPLESVFFSCITTAGSOPVITMTLAAGSL 76
 QY 94 GATLTLTVNFRHVFVAFSPPLAV---KNPVARFSVPALIDEAYAVTAAR---PA 145
 DB 77 WVAALTYMADVDVHYVYGPSLSRSRIIRLOKSKTALM--AFGLTDEVFALAAAKVNRNR 134
 QY 146 GWSAMRLISMOIAHFWYVFGGLTGVAILIPFPEIKGLEFALCSFTLTLDSCRT 201
 DB 135 RSENNMIGIAFSWSWVFGIVIGAFSSGGLGCPANVAAALGMLPALFMSFTLASFQ 194
 QY 202 TKQIPSLLAGISFTIALV-VIPGALFA---ALL-IFGLLT-IRYFGL 247
 DB 195 RKQ-----SLCVTAALVGAALAGVTLFISIVAILAGVCCGLTALIOAFWQ 240

RESULT 12
 Q8FE06 PRELIMINARY; PRT; 245 AA.
 ID Q8FE06;
 AC Q8FE06;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Hypothetical protein ygaZ.
 GN YGAZ OR C3235.

ID Q8G3T5 PRELIMINARY; PRT; 338 AA.
 AC Q8G3T5;
 DT 01-MAR-2003 (TREMBLrel. 23, Created)
 DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
 DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
 DE Narrowly conserved hypothetical protein.
 GN BL669.
 OS Bifidobacterium longum.
 OC Bacteria; Actinobacteria; Actinobacteridae; Bifidobacteriales;
 OC Bifidobacteriaceae; Bifidobacterium.
 OX NCBI_TaxID=216816;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NC 2705;
 RX MEDLINE=2294977; PubMed=12381787;
 RA Schell M.A., Karmali M., Snel B., Vilanova D., Berger B.,
 RA Pessi G., Zwaalen M.-C., Desiere F., Bork P., Delley M.,
 RA Pridmore R.D., Arigoni F.;
 RT "The genome sequence of Bifidobacterium longum reflects its adaptation
 to the human gastrointestinal tract."
 RL Proc. Natl. Acad. Sci. U.S.A. 99:14422-14427(2002).
 DR EMBL; AE014601; MAN25456.1; -
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0006810; P:transport; IEA.
 DR InterPro; IPR004471; AzIC.
 DR Pfam; PF03591; AzIC; 1.
 DR Hypothetical protein; Complete proteome.
 KM SEQUENCE 338 AA; 37538 MW; 805D6F5BA0C08A5 CRC64;
 SQ
 Query Match 16.9%; Score 214; DB 16; Length 338;
 Best Local Similarity 29.9%; Pred. No. 2.6e-09;
 Matches 64; Conservative 37; Mismatches 85; Indels 28; Gaps 8;
 QY 44 LGWYPIGIAFGLVIOGYEWMAAPLFGSLIFAGSTEMVIALVGA-APLGNALTTLL 102
 DB 99 LGFPLDASGYILMGTFGSPFWPMCSAIFPAGSMFVYVLLAFPLNGF-LIALLM 157
 QY 103 VNFHVFYAFSPFLHVNKPIAF-----YVPLIDEAYV--TAAPPGMS-AMR 151
 DB 158 VNARHLFYGLSM-----LGKFKLGMKRPYLFGMCDETFAINSTAKIPAGIDRGWF 209
 QY 152 LISQIAFHSGYVWVGGITGVAINELIPEIKGLFALCSFVLTLDSCRTKQIPSL-L 210
 DB 210 YFWVTLCNQIYVWGTATGLGILGALHPTNDGIDPVLITLFLVIFDQWLDGHRERLSA 269
 QY 211 LAGLSFTIALVPIG-----QALFALLIFLGL 238
 DB 270 VIGVLTSLACLILFGANDFIMPIMIALILFVAL 303
 RESULT 8
 Q8Y223 PRELIMINARY; PRT; 242 AA.
 ID Q8Y223
 AC Q8Y223;
 DT 01-MAR-2002 (TREMBLrel. 20, Created)
 DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
 DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
 DE Hypothetical transmembrane protein RSC0513.
 GN RSC0513 OR RS04992.
 OS Ralstonia solanacearum (Pseudomonas solanacearum).
 OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
 OC Burkholderiaceae; Ralstonia.
 OX NCBI_TaxID=305;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=GM11000;
 RX MEDLINE=21681879; PubMed=11823852;
 RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangelot S.,
 RA Ariat M., Billault A., Broctier P., Camus J.C., Catolico L.,
 RA Chander M., Choise N., Claudel-Renard C., Cunac S., Demange N.,
 RA Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schlex T.,
 RA Sigler P., Trebault P., Whalen M., Wincker P., Levy M.,
 Weissendbach V., Boucher C.A.;

RT "Genome sequence of the plant pathogen Ralstonia solanacearum."
 RL Nature 415:497-502(2002).
 DR EMBL; AL646059; CAD14041.1; -
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0006810; P:transport; IEA.
 DR InterPro; IPR004471; AzIC.
 DR Pfam; PF03591; AzIC; 1.
 KM Hypothetical protein; Complete proteome.
 SQ SEQUENCE 242 AA; 25196 MW; F22503FA454C28B CRC64;
 Query Match 16.5%; Score 209; DB 16; Length 242;
 Best Local Similarity 28.2%; Pred. No. 4.7e-09;
 Matches 57; Conservative 38; Mismatches 95; Indels 12; Gaps 5;
 QY 28 RRYEIAQGLKTSIAGIGWYPIGIAFGLVIOGYEWMAAPLFGSLIFAGSTEMVIALV 87
 DB 5 RAHEFRAGVLTLPMLVGVVPGILGVLTAFAGMPAMILAVANSTIVFGASQMIIVQLM 64
 QY 88 VGAAPLALITLLVNFHVFYAFSPFLHVNKPIA-RYSVPLIDEAYV-----VTA 141
 DB 65 AGGAPALMAATVSMVNLRLALYSASTAPALVHLPRKMLIYVLTDEAFAMNRRVN 124
 QY 142 ARPAGMSA---WELISQIAFHSGYVWVGGITGVAINELIPEIKGLFALCSFVLTLL 197
 DB 125 AQPDAQENTHRHMFPLDAGVALMTSMQASTIVGALGAKVP-ATWPLDFLPLTFIAIV 183
 QY 198 DSCRTKQIPSLIAGLSFTIA 219
 DB 184 PSLTRPQL-APALTGAAALVA 204
 RESULT 9
 Q871S6 PRELIMINARY; PRT; 236 AA.
 ID Q871S6
 AC Q871S6;
 DT 01-JUN-2003 (TREMBLrel. 24, Created)
 DT 01-JUN-2003 (TREMBLrel. 24, Last sequence update)
 DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
 DE AzIC family protein.
 GN VPA0530.
 OS Vibrio parahaemolyticus.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
 OC Vibrionaceae; Vibrio.
 OX NCBI_TaxID=670;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=RIND 221063 / Serotype O3:K6;
 RX MEDLINE=22508454; PubMed=12620739;
 RA Makino K., Oshima K., Kurokawa K., Yokoyama K., Uda T., Tagomori K.,
 RA Iijima Y., Nishida M., Nakano M., Yamashita A., Kubota Y., Kimura S.,
 RA Yasunaga T., Honda T., Shinagawa H., Hattori M., Iida T.;
 RT "Genome sequence of Vibrio parahaemolyticus: a pathogenic mechanism
 distinct from that of V. cholerae."
 RL Lancet 361:743-749(2003).
 DR EMBL; AP005085; EAC61873.1; -
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0006810; P:transport; IEA.
 DR InterPro; IPR004471; AzIC.
 DR Pfam; PF03591; AzIC; 1.
 KM Complete proteome.
 SQ SEQUENCE 236 AA; 25483 MW; 2288BE9F04DB15CB CRC64;
 Query Match 16.0%; Score 203.5; DB 16; Length 236;
 Best Local Similarity 27.1%; Pred. No. 1.3e-08;
 Matches 57; Conservative 44; Mismatches 86; Indels 23; Gaps 8;
 QY 44 LGWYPIGIAFGLVIOGYEWMAAPLFGSLIFAGSTEMVIALVGAAPLGA 96
 DB 21 LAMPSTIAVLWGLLAGSPALDTGLHLEQALAILFAGSAQVAMGMKAGAGLTTM 80
 QY 97 ALTTLLVNFHVFYAFSPFLHVNKPIA-RYSVPLIDEAYVTAAPAG-----WSA 149
 DB 81 LITTFITSRHFLYSVSRKSIAPLPKWRSLIGFLTDELPAI-----AGHSDEQENR 135

RA Springer T.A., Umayam L.A., White O., White R.H., de Macario E.C.,
 RA Ferry J.G., Jarrell K.F., Jing H., Macario A.J.L., Paulsen I.,
 RA Pritchett M., Sowers K.R., Swanson R.V., Zinder S.H., Lander E.,
 RA Metcalf W.W., Birren B.,
 RA "The genome of *Methanosarcina acetivorans* reveals extensive metabolic
 RT and physiological diversity." ;
 RL Genome Res. 12:532-542(2002).
 DR EMBL; AB010501; AA06804.1;
 DR GO; GO:0016021; C: integral to membrane; IEA.
 DR GO; GO:0006810; P: transport; IEA.
 DR InterPro; IPR004471; AzIC.
 DR Pfam; PF03591; AzIC; 1.
 KW Complete proteome.
 SQ SEQUENCE 291 AA; 32067 MW; AFA5100B00B1B CRC64;
 Query Match 23.7%; Score 301; DB 17; Length 291;
 Best Local Similarity 32.6%; Pred. No. 2.4e-16;
 Matches 79; Conservative 38; Mismatches 97; Indels 28; Gaps 6;
 QY 24 DKGRRYRIAG-----LKTSLAAGLGMYPPIGIAFGLVIVQYEWMAAPLPSGLIF 75
 DB 23 EKEVRRYTRGGDRPLFLSALKTTPVFLGYIPLGMAFGFLMDAGYHWIYAFMSLLIY 82
 QY 76 AGSTEMLVIALVGAAPGATLTLVNFHVFYAFSPFLHVKNPPIARF-----Y 127
 DB 83 AGAGCFVALLAAGAGLPEFVIALTLNLHAFYGLSL-----LDKSDIGKVKPY 134
 QY 128 SVEFALIDEAVN--TAAPAGMSANRLISMQIAF-HSYWVFGTLGVAAIEIPFEIKGL 184
 DB 135 LIFALTDERYALLTHVEVPAGSKARFFYISALDHFWAGSVLGAAGSLDNLNMEGM 194
 QY 185 EFALCSLPVTLTDS-CRTKQIPSLIAGLSTIALVIVTGOALFALLIFGLITRY 243
 DB 195 AFVLTALFVLTIEQYFMSIRFPVAAVGAAGASLLIPSENNLLISIIIGTLIARE 254
 QY 244 PF 245
 DB 255 KP 256
 RESULT 5
 Q8B325 PRELIMINARY; PRT; 244 AA.
 AC Q8B325;
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Branched-chain amino acid transport protein.
 GN AZIC OR MM3362.
 OS Methanosarcina mazei (Methanosarcina frisia).
 OC Archaea; Euryarchaeota; Euryarchaeota orders incertae sedis;
 OC Methanosarcinales; Methanosarcinaceae; Methanosarcina.
 OX NCBI_TaxID=2209;
 RX MEDLINE=2210827; PubMed=12125824;
 RP SEQUENCE FROM N.A.
 RP STRAIN=Gel / Gol / ATCC BAA-199 / DSM 3647 / OCM 88;
 RX MEDLINE=2210827; PubMed=12125824;
 RA Depierreux U., Johan A., Hartoch T., Merkl R., Schmitz R.A.,
 RA Martnez-Arias R., Henne A., Wietzer A., Beumer S., Jacob C.,
 RA Bueggemann H., Leneard T., Christmann A., Boemcke M., Steckel S.,
 RA Bhattacharya A., Lykidis A., Overbeek R., Klein H.-P., Gunsalus R.P.,
 RA Fritz H.-J., Gottschalk G.;
 RT "The genome of *Methanosarcina mazei*: evidence for lateral gene
 RT transfer between Bacteria and Archaea." ;
 RT J. Mol. Microbiol. Biotechnol. 4:453-461(2002).
 DR EMBL; AB013585; AA032858.1;
 DR GO; GO:0016021; C: integral to membrane; IEA.
 DR GO; GO:0006810; P: transport; IEA.
 DR InterPro; IPR004471; AzIC.
 DR Pfam; PF03591; AzIC; 1.
 KW Complete proteome.
 SQ SEQUENCE 244 AA; 26655 MW; AC8A74DA99EBCC60 CRC64;

Query Match 22.5%; Score 285; DB 17; Length 244;
 Best Local Similarity 32.8%; Pred. No. 3.9e-15;
 Matches 76; Conservative 40; Mismatches 86; Indels 28; Gaps 9;
 QY 23 DKGRRYRIAG-----LKTSLAAGLGMYPPIGIAFGLVIVQYEWMAAPLPSGLIF 81
 DB 12 DDK-----DVFGLAKTTPVFLGYIPLGMAFGFLMDAGYHWIYAFMSLLVAGSGQF 66
 QY 82 LVIALVGAAPGATLTLVNFHVFYAFSPFLHVKNPPIARF-----Y 133
 DB 67 LAVALLAAGAGLTERVIALTLNLHAFYGLSL-----LEKSDVGKVKPYLIFALT 118
 QY 134 DEAYN--TAAPAGMSANRLISMQIAF-HSYWVFGTLGVAAIEIPFEIKGEFALCS 190
 DB 119 DETYALLTTTEVPKGSRSRFFYIALDHLDTWITGSVLGALLSLDNLNMEGMFVLT 178
 QY 191 LFTVTLTDS-CRTKQIPSLIAGLSTIALVIVTGOALFALLIFGLIT 241
 DB 179 LFTVLTIEQYFMSIRFPVAAVGAAG-TLSLIFPENML-LISIIIGTLIL 228
 RESULT 6
 Q9CNK7 PRELIMINARY; PRT; 240 AA.
 AC Q9CNK7;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Hypothetical protein PM0422.
 GN PM0422.
 OS Pasteurella multocida.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
 OC Pasteurellaceae; Pasteurella.
 OX NCBI_TaxID=747;
 RX MEDLINE=2145866; PubMed=11248100;
 RP SEQUENCE FROM N.A.
 RP STRAIN=Pm70;
 RX MEDLINE=2145866; PubMed=11248100;
 RA May B.J., Zhang Q., Li L.L., Paustian M.L., Whitam T.S., Kapur V.;
 RT "Complete genomic sequence of *Pasteurella multocida* Pm70." ;
 RT Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).
 RL EMBL; AB006078; AA02506.1;
 DR GO; GO:0016021; C: integral to membrane; IEA.
 DR GO; GO:0006810; P: transport; IEA.
 DR InterPro; IPR004471; AzIC.
 DR Pfam; PF03591; AzIC; 1.
 DR TIGRFAMs; TIGR00346; azic; 1.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 240 AA; 26864 MW; 63A1770AC2198BBE CRC64;
 Query Match 18.7%; Score 237; DB 16; Length 240;
 Best Local Similarity 31.3%; Pred. No. 2.7e-11;
 Matches 68; Conservative 41; Mismatches 94; Indels 14; Gaps 8;
 QY 45 GMYPIGIAFGLVIVQYEWMAAPLPSGLIFAGSTEMLVIALVGA-APGATLTLV 103
 DB 26 GFLPVLGIVGVYMGALGEALYPTLMFLYAGSVEIVAGLVAPSPIN-VLLITLMV 84
 QY 104 NFRHVFYAFS-----FPLHVNPNPIARFYSVAFALIDEAVAT--AARPAGMS-ANRLISMQ 156
 DB 85 SGRIQFVIGISMLKXGAVLGRK--RWYLIISTLVDEAFSINVMKVEGIERGVFFVS 141
 QY 157 IAFHSYWFPGTLGVAAIEIPFEIKGEFALCSLFTVTLTDS-CRTKQIPSLIAGLS 215
 DB 142 FYLOFYWVIGAIGALGFSIIIPDLSGIEPMTLFLVIFAEQCKERSHESALIGGIA 201
 QY 216 FTIALVIVPGALFPAALLIFGLITIRYFPL-GXAK 251
 DB 202 FTIALVIVKTFILPILIGIFALTIRVKSATLAK 238
 RESULT 7
 Q8G3T5

Matches 251; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MOKTQEHSSLEVPSPKALPEPDKGYRREYIAQGLKSLAAGLWYPIGIAFGLLVYQ 60
DB 1 MOKTQEHSSLEVPSPKALPEPDKGYRREYIAQGLKSLAAGLWYPIGIAFGLLVYQ 60
QY 61 GYEMWAAPLFSGGLIFASSTEMLVIALVGAAPGALATLTLVNFHVFYAFSPPLHYK 120
DB 61 GYEMWAAPLFSGGLIFASSTEMLVIALVGAAPGALATLTLVNFHVFYAFSPPLHYK 120
QY 121 NPIRAFSVFALIDEAAVAVTARPAAGSAMRLISMOAFHSYVFGGLVGAIAELIPE 180
DB 121 NPIRAFSVFALIDEAAVAVTARPAAGSAMRLISMOAFHSYVFGGLVGAIAELIPE 180
QY 181 IKGLEFALCSLFTVLTLDSCRTKQIPSLLAGISFTIALVPIGQALFALLIFGLLT 240
DB 181 IKGLEFALCSLFTVLTLDSCRTKQIPSLLAGISFTIALVPIGQALFALLIFGLLT 240
QY 241 IRYFFLGKAAK 251
DB 241 IRYFFLGKAAK 251

RESULT 2

Q8FSW8 PRELIMINARY; PRT; 238 AA.

ID Q8FSW8
AC Q8FSW8
RT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DE 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Conserved hypothetical protein.
GN CE2926.
OS Corynebacterium efficiens.
OC Bacteria; Actinobacterias; Actinobacteridae; Actinomycetales;
OC Corynebacteriaceae; Corynebacteriaceae; Corynebacterium.
NCBI_Taxid=152794;

RA Kaveriayasi Y., Yamazaki J., Hino Y., Kikuchi H., Nakamura Y.,
Ikeo K., Suzuki M., Mashima J., Itoh T., Yamagishi A., Niemio Y.,
Usuda Y., Sugimoto S.;
RT "The entire genomic sequence of Corynebacterium efficiens YS-314,"
RL Submitted (MAY-2002) to the EMBL/Genbank/DBJ databases.
DR EMBL; AP005224; BAC19736.1; -
DR GO; GO:0016021; C:Integral to membrane; IEA.
DR GO; GO:0006810; P:Transport; IEA.
DR InterPro; IPR004471; AzLC.
DR Pfam; PF03591; AzLC; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 238 AA; 25739 MW; 02BDEDB0AB33A31 CRC64;

Query Match 35.3%; Score 448.5; DB 16; Length 238;

Best Local Similarity 42.3%; Pred. No. 3e-28; Mismatches 79; Indels 3; Gaps 3;

Matches 91; Conservative 42; Mismatches 79; Indels 3; Gaps 3;
QY 31 EIAQGLKSLAAGLWYPIGIAFGLLVYQGYEMWAAPLFSGGLIFASSTEMLVIALV-VG 89
DB 8 EIAQGLKSLAAGLWYPIGIAFGLLVYQGYEMWAAPLFSGGLIFASSTEMLVIALV-VG 89
QY 8 EIAQGLKSLAAGLWYPIGIAFGLLVYQGYEMWAAPLFSGGLIFASSTEMLVIALV-VG 89
DB 8 EIAQGLKSLAAGLWYPIGIAFGLLVYQGYEMWAAPLFSGGLIFASSTEMLVIALV-VG 89
QY 90 AAPGALATLTLVNFHVFYAFSPPLHYKVPKPIAFYSFALIDEAVATARPAG-WS 148
DB 90 AAPGALATLTLVNFHVFYAFSPPLHYKVPKPIAFYSFALIDEAVATARPAG-WS 148
QY 68 VGPSAL-VTGFVNFRHIFGLTFPRNARSKGRAYSTVALTDAYVIAASRPGEIS 126
DB 68 VGPSAL-VTGFVNFRHIFGLTFPRNARSKGRAYSTVALTDAYVIAASRPGEIS 126
QY 149 AMRLISMOAFHSYVFGGLVGAIAELIPEIKGLEFALCSLFTVLTLDSCRTKQIPS 208
DB 149 AMRLISMOAFHSYVFGGLVGAIAELIPEIKGLEFALCSLFTVLTLDSCRTKQIPS 208
QY 127 GTRLLTQICQALWYPIGIAFGLLVYQGYEMWAAPLFSGGLIFASSTEMLVIALV-VG 186
DB 127 GTRLLTQICQALWYPIGIAFGLLVYQGYEMWAAPLFSGGLIFASSTEMLVIALV-VG 186
QY 209 LLAGLSFTIALVPIGQALFALLIFGLLT 243
DB 209 LLAGLSFTIALVPIGQALFALLIFGLLT 243
QY 187 PLFAVVLAVSGFVAPEQMLVIALTYFILLRVEF 221
DB 187 PLFAVVLAVSGFVAPEQMLVIALTYFILLRVEF 221

RESULT 3

Q8NL66 PRELIMINARY; PRT; 237 AA.

ID Q8NL66
AC Q8NL66
RT 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DE 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Predicted branched-chain amino acid permease (azalucine resistance).
GN CG13083.
OS Corynebacterium glutamicum (Brevibacterium flavum).
OC Bacteria; Actinobacterias; Actinobacteridae; Actinomycetales;
OC Corynebacteriaceae; Corynebacteriaceae; Corynebacterium.
NCBI_Taxid=1718;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;
RA Nakagawa S.;
RT "Complete genomic sequence of Corynebacterium glutamicum ATCC 13032,"
RL Submitted (MAY-2002) to the EMBL/Genbank/DBJ databases.
DR EMBL; AP005283; BAC00477.1; -
DR GO; GO:0016021; C:Integral to membrane; IEA.
DR GO; GO:0006810; P:Transport; IEA.
DR InterPro; IPR004471; AzLC.
DR Pfam; PF03591; AzLC; 1.
KW Complete proteome.
SQ SEQUENCE 237 AA; 25867 MW; 193EB1182E0CF923 CRC64;

Query Match 33.8%; Score 428.5; DB 16; Length 237;

Best Local Similarity 40.1%; Pred. No. 1.2e-26; Mismatches 87; Conservative 46; Mismatches 81; Indels 3; Gaps 3;

QY 31 EIAQGLKSLAAGLWYPIGIAFGLLVYQGYEMWAAPLFSGGLIFASSTEMLVIALV-VG 89
DB 8 EIAQGLKSLAAGLWYPIGIAFGLLVYQGYEMWAAPLFSGGLIFASSTEMLVIALV-VG 89
QY 8 EIAQGLKSLAAGLWYPIGIAFGLLVYQGYEMWAAPLFSGGLIFASSTEMLVIALV-VG 89
DB 8 EIAQGLKSLAAGLWYPIGIAFGLLVYQGYEMWAAPLFSGGLIFASSTEMLVIALV-VG 89
QY 90 AAPGALATLTLVNFHVFYAFSPPLHYKVPKPIAFYSFALIDEAVATARPAG-WS 148
DB 90 AAPGALATLTLVNFHVFYAFSPPLHYKVPKPIAFYSFALIDEAVATARPAG-WS 148
QY 68 IGPSA-AVAGFVNFRHIFGLTFPRNARSKGRAYSTVALTDSEYAVIARPPDIS 126
DB 68 IGPSA-AVAGFVNFRHIFGLTFPRNARSKGRAYSTVALTDSEYAVIARPPDIS 126
QY 149 AMRLISMOAFHSYVFGGLVGAIAELIPEIKGLEFALCSLFTVLTLDSCRTKQIPS 208
DB 149 AMRLISMOAFHSYVFGGLVGAIAELIPEIKGLEFALCSLFTVLTLDSCRTKQIPS 208
QY 127 GTRLLTQICQALWYPIGIAFGLLVYQGYEMWAAPLFSGGLIFASSTEMLVIALV-VG 186
DB 127 GTRLLTQICQALWYPIGIAFGLLVYQGYEMWAAPLFSGGLIFASSTEMLVIALV-VG 186
QY 209 LLAGLSFTIALVPIGQALFALLIFGLLT 245
DB 209 LLAGLSFTIALVPIGQALFALLIFGLLT 245
QY 187 PLFAVVLAVSGFVAPEQMLVIALTYFILLRVEF 223
DB 187 PLFAVVLAVSGFVAPEQMLVIALTYFILLRVEF 223

RESULT 4

Q8TKHO PRELIMINARY; PRT; 291 AA.

ID Q8TKHO
AC Q8TKHO
RT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DE 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Branched chain amino acid transport protein AzLC.
GN MA3437.
OS Methanosarcina acetivorans.
OC Archaea; Euryarchaeota; Euryarchaeota orders incertae sedis;
OC Methanosarcinales; Methanosarcinaceae; Methanosarcina.
NCBI_Taxid=2214;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=C2A / ATCC 35395 / DSM 2834;
RX MEDLINE=21929760; PubMed=11932238;
RA Fitzgerald J.E., Nussbaum C., Roy A., Endrizzi M.G., Macdonald P.,
Allen N., Naylor S., Stange-Thomann N., Deatellano K., Johnson R.,
Linton L., McSwan P., McKernan K., Talamas D., Turrell A., Ye W.,
Zimmer A., Barber R.D., Camm I., Graham D.E., Graham D.A., Guss A.M.,
Heiderich R., Ingram-Smith C., Kuettner H.C., Krzycki J.A.,
Leigh J.A., Li W., Liu J., Mukhopadhyay B., Reeve J.N., Smith K.,

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 16, 2004, 06:51:17 ; Search time 49.6407 seconds
(without alignments)
1595.367 Million cell updates/sec

Title: US-10-608-504-3

Perfect score: 1269

Sequence: 1 MORTGHEHSLSEVSPKAL.....LIFKGLTIRYFKNAK 251

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: SP_ARCHAEA:*
2: SP_BACTERIA:*
3: SP_FUNGI:*
4: SP_HUMAN:*
5: SP_INVERTEBRATE:*
6: SP_MAMMAL:*
7: SP_MHC:*
8: SP_ORGANELLE:*
9: SP_PHAGE:*
10: SP_PLANT:*
11: SP_RODENT:*
12: SP_VIRUS:*
13: SP_VERTEBRATE:*
14: SP_UNCLASSIFIED:*
15: SP_VIRUS:*
16: SP_BACTERIAP:*
17: SP_ARCHAEA:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1269	100.0	251	16	Q8NTP8
2	448.5	35.3	238	16	Q8R5W8
3	428.5	33.8	237	16	Q8N166
4	301	23.7	231	17	Q8TKH0
5	285	22.5	244	17	Q8PS25
6	237	18.7	240	16	Q8CNK7
7	214	16.9	338	16	Q8G3T5
8	209	16.5	242	16	Q8Y223
9	203.5	16.0	236	16	Q8Y1S6
10	200.5	15.8	232	16	Q8YV89
11	198	15.6	245	16	Q8X908
12	198	15.6	245	16	Q8EPO6
13	197	15.5	247	2	Q93KB4
14	194	15.3	230	16	Q88KAS
15	194	15.3	231	16	Q7V198
16	194	15.3	245	16	Q83JZ6

17	192	15.1	250	16	Q8RE08	Q8RE08 fusobacteri
18	186	14.7	237	16	Q8D6C0	Q8D6C0 vibrio vuln
19	185	14.6	230	16	Q8E349	Q8E349 shewanella
20	184.5	14.5	237	16	Q7WU08	Q7WU08 bordetella
21	184.5	14.5	237	16	Q7WB32	Q7WB32 bordetella
22	182	14.3	241	16	Q81SD8	Q81SD8 bacillus an
23	181.5	14.3	251	16	Q81FC0	Q81FC0 bacillus ce
24	181.5	14.3	256	16	Q82R05	Q82R05 rhizobium m
25	181	14.3	234	16	Q8UCR6	Q8UCR6 agrobacteri
26	181	14.3	239	16	Q8KXU8	Q8KXU8 vibrio chol
27	177.5	14.0	218	16	Q8Y1T4	Q8Y1T4 streptococ
28	177	13.9	233	16	Q88Y19	Q88Y19 lactobacill
29	176.5	13.9	235	16	Q83107	Q83107 enterococcu
30	176.5	13.9	242	16	Q88K06	Q88K06 rhizobium 1
31	176	13.8	261	16	Q88T22	Q88T22 lactobacill
32	175	13.8	252	16	Q91276	Q91276 pseudomonas
33	175	13.8	257	16	Q8ZBX0	Q8ZBX0 yersinia pe
34	174.5	13.8	235	16	Q8CF68	Q8CF68 lactococcus
35	174.5	13.8	237	16	Q8K8U2	Q8K8U2 bacillus na
36	169	13.3	235	16	Q8Y761	Q8Y761 listeria mo
37	168	13.2	241	16	Q8CU18	Q8CU18 oceanobacil
38	167	13.2	230	16	Q8E2V7	Q8E2V7 streptococ
39	167	13.2	230	16	Q8DX08	Q8DX08 streptococ
40	166	13.1	235	16	Q92BR4	Q92BR4 listeria in
41	163.5	12.9	219	16	Q8XGA3	Q8XGA3 salmonella
42	163	12.8	238	16	Q82NM2	Q82NM2 rhizobium m
43	161.5	12.7	230	16	Q8CQK2	Q8CQK2 staphylococ
44	161	12.7	224	2	Q8YWS9	Q8YWS9 brucella ab
45	161	12.7	224	16	Q8FYU5	Q8FYU5 brucella su

ALIGNMENTS

RESULT 1

Q8NTP8 PRELIMINARY; PRT; 251 AA.

AC Q8NTP8; 01-OCT-2002 (TrEMBLrel. 22, Created)

DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)

DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)

DE Predicted branched-chain amino acid permease (azalaucine resistance) (Brnf).

GN CG10258 OR BRNF.

OS Corynebacterium glutamicum (Brevibacterium flavum).

OC Bacteria; Actinobacteria; Actinobacteriales; Corynebacteriaceae; Corynebacterium.

OX NCBI_TaxId=1718;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;

RA Nakagawa S.;

RT "Complete genomic sequence of Corynebacterium glutamicum ATCC 13032.";

RL Submitted (MAY-2002) to the EMBL/Genbank/DBJ databases.

RN [2]

RP SEQUENCE FROM N.A.

RX MEDLINE=22077265; PubMed=12081967;

RA Eggeling L.;

RT "Export of L-Isoleucine from Corynebacterium glutamicum: a Two-Gene-Encoded Member of a New Translocase Family.";

RL J. Bacteriol. 184:3947-3956(2002).

DR EMBL; AP005274; BAB97651.1;

DR EMBL; AF454053; AAM46686.1;

DR GO; GO:0016021; C:integral to membrane; IEA.

DR GO; GO:0006810; P:transport; IEA.

DR InterPro; IPR004471; Azic.

DR Pfam; PF03591; Azic; 1.

KW Complete proteome.

SQ SEQUENCE 251 AA; 27333 MW; 898F20D94320D66 CRC64;

Query Match 100.0%; Score 1269; DB 16; Length 251;

Best Local Similarity 100.0%; Pred. No. 6.1e-94;

```

FT VARIANT 67 67 N -> S.
FT VARIANT 102 102 P -> S (IN C/SH2 ANTIGEN).
FT VARIANT 153 153 P -> T (IN RHE KH).
FT VARIANT 225 225 P -> A (IN E/RHS ANTIGEN).
FT VARIANT 232 232 P -> E (IN RHE FM).
FT VARIANT 237 237 P -> V (IN RHE FM).
FT VARIANT 244 244 P -> V (IN VS ANTIGEN).
FT VARIANT 244 244 P -> V (IN RHE FM).
FT CONFLICT 11 11 C -> L (IN REF. 7).
FT CONFLICT 52 52 D -> G (IN REF. 4).
FT CONFLICT 60 60 G -> C (IN REF. 4).
SQ SEQUENCE 416 AA; 45429 MW; DC7AF97DC0DCDF1F CRC64;

Query Match 7.6%; Score 96.5; DB 1; Length 416;
Best Local Similarity 21.7%; Pred. No. 1.3;
Matches 65; Conservative 41; Mismatches 90; Indels 103; Gaps 16;

QY 18 AALEPDDGK-VRRYIAOGLKSLAAGMP-----IGIAGLVIQGYEMWAP 68
DB 35 ASLE-DQKGLVASVGVODLTVMALLGLFLTSNRSSMSVAFNLPMALGVQM--AI 91
QY 69 LPSG-----LIPAGST-----EMVIALVGAAPIGA 95
DB 92 LLDGFLSGPPKGVITLFSRLATMSAVSLISAGAVIKVLAQLVVMVLEVTALGT 151
QY 96 TLL-----TLLVNFRRVFAFS--PELHY--VKMPI-----ARFYSVFALI 133
DB 152 LEMVISNINFTDHYHMLRH-FYFPAVYFGLTVAKCLPELPGRTDNDORATIPSLML 210
QY 134 DEAYVATARPAGMSAMRLISMQI-----AFHSYWGGLTGVAAIELIPEIKG 183
DB 211 GALEFL-----WMFSPVNSPLRSPIQKNAEMNTYY-----ALAVSVTAISG 254
QY 184 LEFALCSLFTVLDSCKTKQIPSLILA-GLSFTIALVVIPOALFALLIFLILTI 241
DB 255 SSIAHPQRKISWT-----YHSAVLAGVAVGTSCHLIPSPMLAVLGLVAGLISI 305

RESULT 15
CLC4 MOUSE STANDARD; PRT; 747 AA.
AC 061418;
AC 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Chloride channel protein 4 (CLC-4).
GN CLC4 OR CLCN4-2 OR CLC4.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SOURCE FROM N.A.
RC STRAIN=CS7BL/6J; TISSUE=Retina;
RX MEDLINE=95400329; PubMed=7670496;
RA Rugarli E.I., Adler D.A., Borsani G., Tsuchiya K., Franco B.,
RA Hauge X., Diatche C., Chapman V., Ballabio A.;
RT "Different chromosomal localization of the Clcn4 gene in Mus spretus
RT and C57BL/6J mice."
RI Nat. Genet. 10:466-471(1995).
CC -!- FUNCTION: Voltage-gated chloride channel. Chloride channels have
CC several functions including the regulation of cell volume;
CC membrane potential stabilization, signal transduction and
CC trans epithelial transport.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- SIMILARITY: Belongs to the chloride channel family.
CC -!- SIMILARITY: Contains 2 CBS domains.

```

```

CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; Z49916; CAA90150.1; -.
DR PIR; I48294; I48294.
DR MGD; MGI:104571; Clcn4-2.
DR InterPro; IPR000644; CBS domain.
DR InterPro; IPR001807; CL-channel_volt.
DR Pfam; PF00571; CBS; 2.
DR Pfam; PF00654; voltage_CLC; 1.
DR PRINTS; PR00762; CHANNEL.
DR SMART; SM00116; CBS; 2.
KW Ionic channel; Ion transport; Chloride channel; Chloride;
KW Voltage-gated channel; Transmembrane; CBS domain; Repeat.
FT TRANSMEM 55 75 POTENTIAL.
FT TRANSMEM 135 155 POTENTIAL.
FT TRANSMEM 189 209 POTENTIAL.
FT TRANSMEM 246 266 POTENTIAL.
FT TRANSMEM 317 337 POTENTIAL.
FT TRANSMEM 352 372 POTENTIAL.
FT TRANSMEM 429 449 POTENTIAL.
FT TRANSMEM 454 474 POTENTIAL.
FT TRANSMEM 495 515 POTENTIAL.
FT TRANSMEM 522 542 POTENTIAL.
FT TRANSMEM 586 646 CBS 1.
FT DOMAIN 684 734 CBS 2.
SQ SEQUENCE 747 AA; 83974 MW; B5486A1B0721144 CRC64;

Query Match 7.5%; Score 95; DB 1; Length 747;
Best Local Similarity 25.6%; Pred. No. 3;
Matches 40; Conservative 25; Mismatches 57; Indels 34; Gaps 8;

QY 80 EMVIALVVG-----AAPLGAIALTLLVNFRRVFAFSPLHVKNPIRFPYSVFALI 134
DB 241 EVLSAAAAGVSAFGAPIGV-----LFSLEVSX--VPELKLTL-----WMSFFPALV 287
QY 135 EAYVATARPAGMSAMRLISMQIAFHSYWGGLTGVAAIELIPEIKG:EPALC-SLFV 193
DB 288 AAFTRISINPFGNS--RLVLFVEYHHPWY-----MDELFFILIGVFGGLNGILFT 337
QY 194 TLTDSCKTKK-----QIPSLLAGLSFTIALVVI 224
DB 338 RCNIAMCRKRKTTLGRYPVLEVAIVAVTAIVAVP 373

```

Search completed: April 16, 2004, 06:53:44
Job time : 16.9833 secs

RN [2] SEQUENCE FROM N.A.
 RP MEDLINE=91058523; PubMed=2123099;
 RX Avent N.D., Ridgwell K., Tanner M.J.A., Anstee D.J.;
 RA "cDNA cloning of a 30 kDa erythrocyte membrane protein associated
 RT with Rh (Rhesus)-blood-group-antigen expression.";
 RL Biochem. J. 271:821-825(1990).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93216282; PubMed=7916743;
 RA Kaji E., Umenishi F., Iwamoto S., Ikemoto S.;
 RT "Isolation of a new cDNA clone encoding an Rh polypeptide associated
 RL with the Rh blood group system.";
 RL Hum. Genet. 91:157-162(1993).
 RN [4]
 RP SEQUENCE FROM N.A. (ISOFORMS RHIV, RHVI AND RHVII).
 RC TISSUE=Bone marrow;
 RX MEDLINE=92360855; PubMed=1379850;
 RA le van Kim C., Cherif-zahar B., Raynal V., Mouro I., Lopez M.,
 RA Cartzon J.-P., Colin Y.;
 RT "Multiple Rh messenger RNA isoforms are produced by alternative
 RL splicing.";
 RL Blood 80:1074-1078(1992).
 RN [5]
 RP SEQUENCE FROM N.A., AND VARIANTS THR-153; GLU-232 AND VAL-237.
 RX MEDLINE=89134163; PubMed=3146980;
 RA Avent N.D., Ridgwell K., Mawby W.J., Tanner M.J.A., Anstee D.J.,
 RA Kumpel B.;
 RT "Protein-sequence studies on Rh-related polypeptides suggest the
 RT presence of at least two groups of proteins which associate in the
 RL human red-cell membrane.";
 RL Biochem. J. 256:1043-1046(1988).
 RN [7]
 RP SEQUENCE OF 1-16.
 RX MEDLINE=88294325; PubMed=3135863;
 RA Bloy C., Blanchard D., Dahr W., Beyreuther K., Salmon C.,
 RA Cartzon J.-P.;
 RT "Determination of the N-terminal sequence of human red cell Rh(D)
 RT polypeptide and demonstration that the Rh(D), (C), and (E) antigens
 RL are carried by distinct polypeptide chains.";
 RL Blood 72:661-666(1988).
 RN [8]
 RP SEQUENCE OF 1-10 FROM N.A.
 RX MEDLINE=9445182; PubMed=8188244;
 RA Cherif-zahar B., le van Kim C., Rouillac C., Raynal V., Cartzon J.-P.,
 RA Colin Y.;
 RT "Organization of the gene (RHE) encoding the human blood group
 RT RhCBE antigens and characterization of the promoter region.";
 RL Genomics 15:68-74(1994).
 RN [9]
 RP VARIANTS BLOOD GROUP C AND E.
 RX MEDLINE=94035121; PubMed=8220426;
 RA Mouro I., Colin Y., Cherif-zahar B., Cartzon J.-P., le van Kim C.;
 RT "Molecular genetic basis of the human Rhesus blood group system.";
 RL Nat. Genet. 5:62-65(1993).
 CC -1- FUNCTION: MAY BE PART OF AN OLIGOMERIC COMPLEX WHICH IS LIKELY TO
 CC SUBCELLULAR LOCATION: CHANNEL FUNCTION IN THE ERYTHROCYTE MEMBRANE.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- ALTERNATIVE PRODUCTS: Integral membrane protein.
 CC Event=Alternative splicing; Named isoforms=4;
 CC Name=RHI;
 CC IsoId=PI8577-1; Sequence=Displayed;
 CC Name=RHIV;
 CC IsoId=PI8577-2; Sequence=VSP_005703, VSP_005704;
 CC Name=RHVI;
 CC IsoId=PI8577-3; Sequence=VSP_005702, VSP_005705;
 CC Name=RHVII;
 CC

```
CC      IsoId=I18577-4; Sequence=VSP_005701;  
CC      TISSUE SPECIFICITY: RESTRICTED TO TISSUES OR CELL LINES EXPRESSING  
CC      ERYTHROID CHARACTERS.  
CC      -| POLYMORPHISM: RhCE and Rhd are responsible for the RH blood group  
CC      system. The molecular basis of the B-RhJ/=Rh5 blood group  
CC      antigens is a single variation in position 225; Pro-225  
CC      corresponds to Rhj and Ala-225 to Rh5. The molecular basis of the  
CC      CsnH/c-nH4 blood group antigens is a single variation in position  
CC      102; Ser-102 corresponds to Rh2 and Pro-102 to Rh4.  
CC      -| SIMILARITY: BELONGS TO THE RH FAMILY. STRONG, TO RHD.  
-----  
CC      This Swiss-Prot entry is copyright. It is produced through a collaboration  
CC      between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC      the European Bioinformatics Institute. There are no restrictions on its  
CC      use by non-profit institutions as long as its content is in no way  
CC      modified and this statement is not removed. Usage by and for commercial  
CC      entities requires a license agreement (See http://www.isb.ch/announce/  
CC      or send an email to licenses@isb.eib.ch).  
-----  
DR      EMBL; M34015; AAA36567.1; +  
DR      EMBL; X54534; CAA38401.1; -  
DR      EMBL; S57967; AAB26080.1; -  
DR      EMBL; X63095; CAA44809.1; +  
DR      EMBL; X63096; CAA44810.1; -  
DR      EMBL; X63098; CAA44812.1; -  
DR      EMBL; AB018644; BAAs3927.1; -  
DR      EMBL; AB018645; BAAs3928.1; -  
DR      EMBL; S70456; AAD14061.1; +  
DR      PIR; A30405; A30405.  
DR      Genew; HGNC:10008; RHCE.  
DR      MIM; 111690; -.  
DR      GO; GO:0005887; C:integral to plasma membrane; TAS.  
DR      InterPro; IPR001905; Ammonium_transp.  
DR      InterPro; IPR002229; RheusRHD.  
DR      Pfam; PF00909; Ammonium_transp; 1.  
DR      PRINTS; PR00342; RHESUSRHD.  
KW      Erythrocytes; Transmembrane; Blood group antigen; Alternative splicing;  
KW      Polymorphism.  
FT      INIT_MET          0  
FT      TRANSMEM         11       31     POTENTIAL.  
FT      TRANSMEM         43       63     POTENTIAL.  
FT      TRANSMEM         76       96     POTENTIAL.  
FT      TRANSMEM        124      144     POTENTIAL.  
FT      TRANSMEM        171      191     POTENTIAL.  
FT      TRANSMEM        202      222     POTENTIAL.  
FT      TRANSMEM        237      257     POTENTIAL.  
FT      TRANSMEM        264      284     POTENTIAL.  
FT      TRANSMEM        286      306     POTENTIAL.  
FT      TRANSMEM        330      350     POTENTIAL.  
FT      TRANSMEM        357      377     POTENTIAL.  
FT      VARSPPLIC        162      312     Missing (in isoform RHVII).  
FT      VARSPPLIC        163      267     Missing (in isoform RHVI).  
FT      FTID=VARPPLIC        313      353     /FTid=VSP_005702.  
FT      VCCNRLVLGIHHISVMHSITSLGLIGETIIYLLLVLTHTWN  
FT      -> DWLPDPQHMGTQLGHRDSSHWSPRPAPKSNMEST  
FT      SCG (in isoform RHIV) .  
FT      /FTid=VSP_005703.  
FT      Missing (in isoform RHIV) .  
FT      /FTid=VSP_005704.  
FT      ISOFORM RHVI)  
FT      KTFDDGVFWKFHLAVGF -> FAPKSNNMSTSC (in  
FT      isoform RHVI) .  
FT      /FTid=VSP_005705.  
FT      W -> C .  
FT      /FTid=VAR_006911.  
FT      A -> T (IN C(X)/RH9 ANTIGEN) .  
FT      /FTid=VAR_006912.  
FT      Q -> R (IN C(W)/RH8 ANTIGEN) .  
FT      /FTid=VAR_006913.  
FT      L -> I .  
FT      /FTid=VAR_006914.
```

```

OY 77 GSTEMVLVAVG-----VGAAPLGAIALTLTLV--NFRHVFYAFSPFLHVNNPIARFYS 128
DB 237 GAVNTAAVVLVPDPPWIGAMWGLISSALVLSFPPFPFPMPIGAKAP----- 290
OY 129 VPAIDEAAYVTAAPAGWSAMWELISMQIAFHSYVWFGGLTGAVALIELIP-PEIKGLEPA 187
DB 291 --ATADBAKLEBAKSRG-----SLVDFPKFPCIFLRLL 323
OY 188 LGSIFVTLTLDSCRTKKQIPSLIAGLS-----FITALVVIPOGL 228
DB 324 MWSLFLVLAQCTF-----SVIAGLSIFLAKFLKXGYGTSAYANFLIGAVNIPAAAL 378
OY 229 FALLIFLGLTLTRYPF 245
DB 379 G---MFGGILMRFFVF 392

RESULT 13
CLC4_HUMAN STANDARD; PRT; 760 AA.
AC P51793; Q9UBU1;
DT 01-OCT-1996 (Rel. 34, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Chloride channel protein 4 (CLC-4).
GN CLC4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Retina;
RA MEDLINE=94348498; PubMed=8069296;
RX van Slegtenhorst M.A., Bassi M.T., Borsani G., Wapenaar M.C.,
RA Ferrero G.B., de Conchitis L., Rugari E.I., Grillo A., Franco B.,
RA Zoghbi H.Y., Ballabio A.;
RT "A gene from the Xp22.3 region shares homology with voltage-gated
RT chloride channels."
RT Hum. Mol. Gene. 3:547-552 (1994).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Skeletal muscle;
RX MEDLINE=20035030; PubMed=10564087;
RA Kawasaki M., Fukuma T., Yamachi K., Sakamoto H., Marumo F.,
RA Sasaki S.;
RT "Identification of an acid-activated Cl- channel from human skeletal
RT muscles."
RT Am. J. Physiol. 277:C948-C954 (1999).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Lens epithelium;
RA "A chloride channel (CLC-4) in human lens epithelium."
RT Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Voltage-gated chloride channel. Chloride channels have
CC several functions including the regulation of cell volume;
CC membrane potential stabilization, signal transduction and
CC transepithelial transport.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- TISSUE SPECIFICITY: Abundant in skeletal muscle and also
CC detectable in brain and heart.
CC -1- SIMILARITY: Belongs to the chloride channel family.
CC -1- SIMILARITY: Contains 2 CBS domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL consortium -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----

```

```

DR EMBL; X77197; CA54417.1; -.
DR EMBL; AB019432; BAA77327.1; -.
DR EMBL; AF170492; AAD50981.1; -.
DR PIR; I37242; I37242.
DR Genew; H8NC; 2022; CLC4.
DR MIM; 302910; -.
DR GO; GO:0005247; F: voltage-gated chloride channel activity; TAS.
DR GO; GO:0006810; P: transport; TAS.
DR InterPro; IPR000644; CBS domain.
DR InterPro; IPR001807; Cl-Channel_volt.
DR Pfam; PF00571; CBS; 2. Channel_volt.
DR Pfam; PF00654; voltage_CLC; 1.
DR SMART; SM00116; CBS; 2.
DR PRINTS; PR00762; CLCHANNEL.
DR SMART; SM00116; CBS; 2.
KW Ionic channel; Ion transport; Chloride channel; Chloride;
KW Voltage-gated channel; Transmembrane; CBS domain; Repeat.
FT TRANSMEM 68 88 POTENTIAL.
FT TRANSMEM 148 168 POTENTIAL.
FT TRANSMEM 202 222 POTENTIAL.
FT TRANSMEM 259 279 POTENTIAL.
FT TRANSMEM 330 350 POTENTIAL.
FT TRANSMEM 365 385 POTENTIAL.
FT TRANSMEM 442 462 POTENTIAL.
FT TRANSMEM 467 487 POTENTIAL.
FT TRANSMEM 508 528 POTENTIAL.
FT TRANSMEM 535 555 POTENTIAL.
FT TRANSMEM 599 659 CBS 1.
FT DOMAIN 697 747 CBS 2.
FT DOMAIN 178 178 A -> R (IN REF. 1).
FT CONFLICT 178 178 II -> YY (IN REF. 1).
FT CONFLICT 498 499 K -> N (IN REF. 1).
FT CONFLICT 659 659
SQ SEQUENCE 760 AA; 84916 MM; 3A5A25DLFEFF217 CRC64;

Query Match 7.7%; Score 98; DB 1; Length 760;
Best Local Similarity 25.6%; Pred. No. 1.8;
Matches 40; Conservative 27; Mismatches 55; Indels 34; Gaps 8;

OY 80 EMLVIALVVG-----AAPLGAIALTLTLVNRHVFYAFSPFLHVNNPIARFYSVPLID 134
DB 254 EVLSAAAAGVSVFAGPAGIGV-----LFSLEEVSY--YFPLKTL-----WSPFALV 300
OY 135 EAYAVTAAPAGWSAMWELISMQIAFHSYVWFGGLTGAVALIELIPPEIKGLEFALC-SLFLV 193
DB 301 AAFILRSINPGNS--RLVLFYVEYHTPMY-----MDELFFILLGVFGGLMGTLLFI 350
OY 194 TLITDSCRTK-----QIPSLIAGLSFTALVVI 224
DB 351 RCNIAMCRRKRTTLGRYPVLEIVVTAITAIIVP 386

RESULT 14
RCCE_HUMAN STANDARD; PRT; 416 AA.
AC P18577; Q02163; Q02164; Q02165; Q16160; Q9UEC2; Q9UEC3;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Blood group Rh (CB) polypeptide (Rhesus C/E antigens) (RH30A) (RH19X)
DE (Rh polypeptide 1) (Rh1).
DE RHCE OR RHC OR RHE.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A. AND PARTIAL SEQUENCE.
RC TISSUE=Bone marrow;
RX MEDLINE=90349591; PubMed=1696722;
RA Cherif-Zahar B., Bloy C., de Van Kim C., Blanchard D., Bailly P.,
RA Hermand P., Salmon C., Carton J.-P., Colin Y.;
RT "Molecular cloning and protein structure of a human blood group Rh
RT polypeptide."
RT Proc. Natl. Acad. Sci. U.S.A. 87:6243-6247 (1990).

```

```

CC or send an email to license@sb-sib.ch).
CC -----
DR EMBL; Z56277; CA91216.1; -.
DR EMBL; D50497; BA09091.1; -.
DR InterPro; IPR000644; CBS domain.
DR InterPro; IPR001807; Cl-channel_volt.
DR Pfam; PF00571; CBS; 2.
DR Pfam; PF00654; voltage CIC; 1.
DR PRINTS; PR00762; CLCHANNEL.
DR SMART; SM00116; CBS; 2.
KW Ionic channel; ion transport; Chloride channel; Chloride;
KV Voltage-gated channel; Transmembrane; CBS domain; Repeat.
FT TRANSMEM 55 75 POTENTIAL.
FT TRANSMEM 133 153 POTENTIAL.
FT TRANSMEM 189 209 POTENTIAL.
FT TRANSMEM 246 266 POTENTIAL.
FT TRANSMEM 317 337 POTENTIAL.
FT TRANSMEM 352 372 POTENTIAL.
FT TRANSMEM 428 448 POTENTIAL.
FT TRANSMEM 453 473 POTENTIAL.
FT TRANSMEM 484 514 POTENTIAL.
FT TRANSMEM 521 541 POTENTIAL.
FT DOMAIN 585 614 CBS 1.
FT DOMAIN 662 733 CBS 2.
FT CONFLICT 315 315 H -> Y (IN REF. 2).
SQ SEQUENCE 746 AA; 83067 MW; 5F17D45F937003CE CRC64;

Query Match 7.9%; Score 100.5; DB 1; Length 746;
Best Local Similarity 22.1%; Pred. No. 1.1;
Matches 50; Conservative 33; Mismatches 68; Indels 75; Gaps 10;

27 YRRYEIAQGLKTSIAAGLGMVPIGIAFGLLVIQGYEMWADPLFGLIFAGSTEMLVITAL 86
232 YRNEKREKREVLASAAAAG--VSVAFG----- 256
DB
QY 87 VCAALPLGALITLTLVNRHRYAPSPPLHVVKNPIAFVSVPALLIDEAVALTARPAQ 146
257 ---ADIGV---LFSLEVSY--VFPLKTL---WSFFPALVAALFTRGINSFPG 299
DB
QY 147 WSAWRLISMQIAFHSYVFGGLGVAIAELIPFRIKGLFALC-SIFVTLTLDSCKTK- 204
300 NS--RLVLFVFEHFTW-----HLFELVPIVIGIRGALFIRTNIAMCRKXKT 349
DB
QY 205 ---QIPSLIAGSTIALVTPGQ-----ALFALLIFLGL 239
350 TOLGKTPVEVVLIVTALTAIPNEYTKMSTSELSELFDGGL 395
DB

RESULT 12
S212_HUMAN STANDARD; PRT; 643 AA.
ID S212_HUMAN
AC Q92939;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Solute carrier family 21 member 2 (Prostaglandin transporter) (PGT).
GN SLC21A2
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OC NCBI_TaxID=9606;
OX 1;
RN SEQUENCE FROM N.A.
RP MEDLINE=96379664; PubMed=8787677;
RA Lu R., Kanai N., Bao Y., Schuster V.L.;
RT "Cloning, in vitro expression, and tissue distribution of a human
RT prostaglandin transporter cDNA(hPGT).";
RL J. Clin. Invest. 98:1142-1149(1996).
RN 12;
RP SEQUENCE FROM N.A.
RX MEDLINE=98289612; PubMed=9618293;
RA Lu R., Schuster V.L.;
RT "Molecular cloning of the gene for the human prostaglandin transporter

```

```

RT hPGT: gene organization, promoter activity, and chromosomal
RT localization."
RU Biochem. Biophys. Res. Commun. 246:805-812(1998).
CC -1- FUNCTION: May mediate the release of newly synthesized
CC prostaglandins from cells, the transepithelial transport of
CC prostaglandins, and the clearance of prostaglandins from the
CC circulation. Transports PGD2, as well as PGE1, PGE2 and PGF2A.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- TISSUE SPECIFICITY: Ubiquitous.
CC -1- SIMILARITY: BELONGS TO THE SLC21 FAMILY OF TRANSPORTERS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.sib.ch/announce/
CC or send an email to license@sb-sib.ch).
CC -----
DR EMBL; U70867; AAC9469.1; -.
DR EMBL; AF056732; AAC62004.1; JOINED.
DR EMBL; AF056719; AAC62004.1; JOINED.
DR EMBL; AF056720; AAC62004.1; JOINED.
DR EMBL; AF056721; AAC62004.1; JOINED.
DR EMBL; AF056722; AAC62004.1; JOINED.
DR EMBL; AF056723; AAC62004.1; JOINED.
DR EMBL; AF056724; AAC62004.1; JOINED.
DR EMBL; AF056725; AAC62004.1; JOINED.
DR EMBL; AF056726; AAC62004.1; JOINED.
DR EMBL; AF056727; AAC62004.1; JOINED.
DR EMBL; AF056728; AAC62004.1; JOINED.
DR EMBL; AF056729; AAC62004.1; JOINED.
DR EMBL; AF056730; AAC62004.1; JOINED.
DR EMBL; AF056731; AAC62004.1; JOINED.
DR EMBL; HGNC:10955; SLC21A2.
DR MIM; 601460; -.
DR GO; GO:0005887; C:integral to plasma membrane; TAS.
DR GO; GO:0005624; C:membrane fraction; TAS.
DR GO; GO:0005319; F:lipid transporter activity; TAS.
DR GO; GO:0006869; P:lipid transport; TAS.
DR InterPro; IPR007114; MFS.
DR InterPro; IPR004157; CATP_Cterm.
DR InterPro; IPR004156; CATP_Nterm.
DR Pfam; PF03137; CATP_C; 1.
DR Pfam; PF03132; CATP_N; 1.
DR TIGRfam; TIGR00805; cat; 1.
DR PROSITE; PS50850; MFS; 1.
DR TRANSMEMBRANE; Transport; Glycoprotein.
KW TRANSMEM 27 47 POTENTIAL.
FT TRANSMEM 71 91 POTENTIAL.
FT TRANSMEM 101 121 POTENTIAL.
FT TRANSMEM 172 192 POTENTIAL.
FT TRANSMEM 210 230 POTENTIAL.
FT TRANSMEM 258 278 POTENTIAL.
FT TRANSMEM 328 348 POTENTIAL.
FT TRANSMEM 366 386 POTENTIAL.
FT TRANSMEM 400 420 POTENTIAL.
FT TRANSMEM 440 460 POTENTIAL.
FT TRANSMEM 511 531 POTENTIAL.
FT TRANSMEM 551 571 POTENTIAL.
FT TRANSMEM 606 626 POTENTIAL.
FT CARBOHYD 134 134 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 478 478 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 491 491 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 643 AA; 70116 MW; 7FC434A53DE2C9E CRC64;

Query Match 7.8%; Score 98.5; DB 1; Length 643;
Best Local Similarity 22.6%; Pred. No. 1.4;
Matches 58; Conservative 30; Mismatches 76; Indels 93; Gaps 10;

QY 40 LAAIGMTPIGIAFGLVYQGYEMWADPLFGLIFAGSTEMLVITAL 76
DB 178 LAAIGMTPIGIAFGLVYQGYEMWADPLFGLIFAGSTEMLVITAL 236

```

```

Db      232 YRKNAKREVLASAAAAG--VSAVFG-----256
Qy      87 VGGAAPLGAIALTTLLVNFPHVFAFSPFLHYVKNPIAFYSVPALIDEAVATAAPAG 146
Db      257 ----APIGV-----LFSLEVSY--YFPLKTL-----WRSFFAALVAAFILRSINPFG 299
Qy      147 MSAMRLISMQIAFHSYVFGSLTGVAIAELIPEIKGLEFALC-SLFTVTLTDSGCRTK- 204
Db      300 NS--RLVLFYVEFHTPW-----HLFELVPEIVIGIFGGLWGALFIRTNIAWCRKRT 349
Qy      205 ----QIPSLLAGISFTIALVVIPOG-----ALFALLIFGLL 239
Db      350 TOLGKYPVEVLIVTATITAILAFNRYTRMSTSELISELFDGGL 395

RESULT 10
ID      CLC5_MOUSE      STANDARD;      PRT;      746 AA.
AC      09WVDA:
DT      16-OCT-2001 (Rel. 40, Created)
DT      10-OCT-2003 (Rel. 42, Last sequence update)
DE      Chloride channel protein 5 (CLC-5).
GN      CLC5 OR CLC5.
OS      Mus musculus (Mouse).
OC      Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX      NCBI_Taxid=10090;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      TISSUE=Kidney;
RX      MEDLINE=99303559; PubMed=10373326;
RA      Tanka K., Fisher S.E., Craig I.W.;
RT      "Characterization of novel promoter and enhancer elements of the mouse
RT      homologue of the Dent disease gene, CLCNS, implicated in X-linked
RT      hereditary nephrolithiasis.";
RL      Genomics 58:281-292(1999).
CC      -1- FUNCTION: Voltage-gated chloride channel. Chloride channels have
CC      several functions including the regulation of cell volume;
CC      membrane potential stabilization, signal transduction and
CC      transepithelial transport. May play an important role in renal
CC      tubular function.
CC      -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC      -1- TISSUE SPECIFICITY: Kidney specific.
CC      -1- SIMILARITY: Belongs to the chloride channel family.
CC      -1- SIMILARITY: Contains 2 CBS domains.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL: AF134117; AAD28473.1; -.
DR      MGD; WGI:99486; CLCNS.
DR      InterPro; IPR000644; CBS_domain.
DR      InterPro; IPR001807; Cl-channel_volt.
DR      Pfam; PF00571; CBS; 2.
DR      Pfam; PF00654; voltage_CLC; 1.
DR      PRINTS; PR00762; CICHANNEL.
DR      SMART; SM00116; CBS; 2.
KW      Ionic channel; Ion transport; Chloride channel; Chloride;
KW      Voltage-gated channel; Transmembrane; CBS domain; Repeat.
FT      TRANSMEM      55
FT      TRANSMEM      75
FT      TRANSMEM      133
FT      TRANSMEM      153
FT      TRANSMEM      189
FT      TRANSMEM      209
FT      TRANSMEM      246
FT      TRANSMEM      266
FT      TRANSMEM      317
FT      TRANSMEM      337
FT      TRANSMEM      352
FT      TRANSMEM      372
FT      TRANSMEM      428
FT      TRANSMEM      448
FT      TRANSMEM      POTENTIAL.

```

```

FT      TRANSMEM      453
FT      TRANSMEM      473
FT      TRANSMEM      494
FT      TRANSMEM      514
FT      TRANSMEM      521
FT      TRANSMEM      541
FT      TRANSMEM      561
FT      TRANSMEM      582
FT      TRANSMEM      614
FT      TRANSMEM      682
FT      TRANSMEM      733
FT      TRANSMEM      746 AA; 83100 MW; DBF3AE4FC331A08 CRC64;
SQ      SEQUENCE

Query Match      7.9%; Score 100.5; DB 1; Length 746;
Best Local Similarity 22.1%; Pred. No. 1.1;
Matches 50; Conservative 33; Mismatches 68; Indels 75; Gaps 10;

Qy      27 YRRYEIAQGLKTSIAGLGMVPIGIAFGGLVIOGYEMAAPLPSGLIFAGSTEMLVIAL 86
Db      232 YRKNAKREVLASAAAAG--VSAVFG-----256
Qy      87 VGGAAPLGAIALTTLLVNFPHVFAFSPFLHYVKNPIAFYSVPALIDEAVATAAPAG 146
Db      257 ----APIGV-----LFSLEVSY--YFPLKTL-----WRSFFAALVAAFILRSINPFG 299
Qy      147 MSAMRLISMQIAFHSYVFGSLTGVAIAELIPEIKGLEFALC-SLFTVTLTDSGCRTK- 204
Db      300 NS--RLVLFYVEFHTPW-----HLFELVPEIVIGIFGGLWGALFIRTNIAWCRKRT 349
Qy      205 ----QIPSLLAGISFTIALVVIPOG-----ALFALLIFGLL 239
Db      350 TOLGKYPVEVLIVTATITAILAFNRYTRMSTSELISELFDGGL 395

RESULT 11
ID      CLC5_RAT      STANDARD;      PRT;      746 AA.
AC      P51796; P70642;
DT      01-OCT-1996 (Rel. 34, Created)
DT      01-OCT-1996 (Rel. 34, Last sequence update)
DE      "Cloning and functional expression of rat CLC-5, a chloride channel
DE      homologue of the Dent disease gene."
GN      CLCNS.
OS      Rattus norvegicus (Rat).
OC      Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX      NCBI_Taxid=10116;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      TISSUE=Brain;
RX      MEDLINE=96125100; PubMed=8537381;
RA      Steinmeyer K., Schwappach B., Bens M., Vandewalle A., Dentsch T.U.;
RT      "Cloning and functional expression of rat CLC-5, a chloride channel
RT      related to kidney disease.";
RL      J. Biol. Chem. 270:31172-31177(1995).
RN      [2]
RP      SEQUENCE FROM N.A.
RC      TISSUE=Kidney;
RX      MEDLINE=96215316; PubMed=8626585;
RA      Sakamoto H., Kawasaki M., Uchida S., Sasaki S., Maruno F.;
RT      "Identification of a new outwardly rectifying Cl- channel that
RT      belongs to a subfamily of the CLC Cl- channels.";
RL      J. Biol. Chem. 271:10210-10216(1996).
CC      -1- FUNCTION: Voltage-gated chloride channel. Chloride channels have
CC      several functions including the regulation of cell volume;
CC      membrane potential stabilization, signal transduction and
CC      transepithelial transport. May play an important role in renal
CC      tubular function.
CC      -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC      -1- TISSUE SPECIFICITY: Kidney specific.
CC      -1- SIMILARITY: Belongs to the chloride channel family.
CC      -1- SIMILARITY: Contains 2 CBS domains.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/

```

GN CLCN5 OR CLCN2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OC NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Kidney;
 RX MEDLINE=96121370; PubMed=8575751;
 RA Fisher S.E., van Bakel I., Lloyd S.E., Pearce S.H.S.,
 RA Thakker R.V., Craig I.W.;
 RT "Cloning and characterization of CLCN5, the human kidney chloride
 channel gene implicated in Dent disease (an X-linked hereditary
 nephrolithiasis)".
 RL Genomics 29:598-606(1995).
 RN [2]
 RP SEQUENCE OF 487-746 FROM N.A.
 RC TISSUE=Kidney;
 RX MEDLINE=95179126; PubMed=7874126;
 RA Fisher S., Black G.C.M., Lloyd S.E., Hatchwell E., Wrong O.,
 RA Thakker R.V., Craig I.W.;
 RT "Isolation and partial characterization of a chloride channel gene
 which is expressed in kidney and is a candidate for Dent's disease
 (an X-linked hereditary nephrolithiasis)".
 RL Hum. Mol. Genet. 3:2053-2059(1994).
 RN [3]
 RP TISSUE SPECIFICITY.
 RC TISSUE=Aortic endothelium, and Vascular smooth muscle;
 RX MEDLINE=9922497; PubMed=10198195;
 RA Lamb P.S., Clayton G.H., Liu B.-X., Smith R.L., Barna T.J.,
 RA Schulte B.C.;
 RT "Expression of CLCN voltage-gated chloride channel genes in human
 blood vessels".
 RL J. Mol. Cell. Cardiol. 31:657-666(1999).
 RN [4]
 RP VARIANTS NPHL ARG-200; GLU-506 AND PRO-520, AND VARIANT XLRH LEU-244.
 RX MEDLINE=96158876; PubMed=8559248;
 RA Lloyd S.E., Pearce S.H.S., Fisher S.E., Steimeyer K., Schwappach B.,
 RA Scheinman S.J., Harding B., Bolino A., Devoto M., Goodyer P.,
 RA Rigidin S.P.A., Wrong O., Jentsch T.J., Craig I.W., Thakker R.V.;
 RT "A common molecular basis for three inherited kidney stone diseases".
 RL Nature 379:445-449(1996).
 RN [5]
 RP VARIANTS NPHL.
 RX MEDLINE=97402204; PubMed=9259268;
 RA Lloyd S.E., Guenther W., Pearce S.H.S., Thomson A., Bianchi M.L.,
 RA Bosio M., Craig I.W., Fisher S.E., Scheinman S.J., Wrong O.,
 RA Jentsch T.J., Thakker R.V.;
 RT "Characterization of renal chloride channel, CLCN5, mutations in
 hypercalcaemic nephrolithiasis (kidney stones) disorders".
 RL Hum. Mol. Genet. 6:1233-1239(1997).
 CC -1- FUNCTION: Voltage-gated chloride channel. Chloride channels have
 several functions including the regulation of cell volume,
 membrane potential stabilization, signal transduction and
 transepithelial transport. May play an important role in renal
 tubular function.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- TISSUE SPECIFICITY: Kidney. Moderately expressed in aortic
 vascular smooth muscle and endothelial cells, and at a slightly
 higher level in the coronary vascular smooth muscle.
 CC -1- DISEASE: Defects in CLCN5 are a cause of X-linked recessive
 hypophosphataemic rickets (XLRH) [MIM:300008].
 CC -1- DISEASE: Defects in CLCN5 are the cause of idiopathic low
 molecular weight proteinuria of Japanese children (JILP)
 [MIM:300009].
 CC -1- DISEASE: Defects in CLCN5 are the cause of X-linked recessive
 nephrolithiasis type 2 (NPHL2) [MIM:300099], also known as Dent
 disease. NPHL2 is a renal form of Fanconi syndrome, characterized
 by hypercalcaemia, nephrocalcinosis, renal stones and progressive
 renal failure.
 CC -1- DISEASE: Defects in CLCN5 are the cause of X-linked recessive
 nephrolithiasis type 1 (NPHL1) [MIM:310468]; also designated XRN.
 CC NPHL1 is characterized by hypercalcaemia, nephrocalcinosis, renal

CC stones and renal insufficiency, but differs from Dent disease in
 CC that the patients lack urinary acidification defects, rickets, and
 CC osteomalacia.
 CC -1- SIMILARITY: Contains 2 CBS domains.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X91906; CA63000.1; -;
 DR EMBL; X81836; CA57430.1; -;
 DR PIR; I37277; I37277.
 DR GeneW; HGNC:2023; CLCN5.
 DR MIM; 300008; -;
 DR MIM; 300009; -;
 DR MIM; 310468; -;
 DR GO; GO:0005887; C:integral to plasma membrane; TAS.
 DR GO; GO:0005254; F:chloride channel activity; TAS.
 DR GO; GO:0007586; P:excretion; TAS.
 DR GO; GO:0006810; P:transport; TAS.
 DR InterPro; IPR000644; CBS_domain.
 DR InterPro; IPR01807; Cl_Channel_volt.
 DR Pfam; PF00571; CBS; 2.
 DR Pfam; PF00654; voltage_CLC; 1.
 DR PRINTS; PR00762; CLCHANNEL.
 DR SMART; SM0116; CBS; 2.
 KW Ionic channel; Ion transport; Chloride channel; Chloride;
 KW Voltage-gated channel; Transmembrane; CBS domain; Repeat;
 KW Disease mutation.
 FT TRANSSEM 55 75 POTENTIAL.
 FT TRANSSEM 133 153 POTENTIAL.
 FT TRANSSEM 189 209 POTENTIAL.
 FT TRANSSEM 246 266 POTENTIAL.
 FT TRANSSEM 317 337 POTENTIAL.
 FT TRANSSEM 352 372 POTENTIAL.
 FT TRANSSEM 428 448 POTENTIAL.
 FT TRANSSEM 453 473 POTENTIAL.
 FT TRANSSEM 494 514 POTENTIAL.
 FT TRANSSEM 521 541 POTENTIAL.
 FT DOMAIN 585 645 CBS 1.
 FT DOMAIN 682 733 CBS 2.
 FT VARIANT 30 30
 FT VARIANT 57 57
 FT VARIANT 200 200
 FT VARIANT 244 244
 FT VARIANT 280 280
 FT VARIANT 506 506
 FT VARIANT 512 512
 FT VARIANT 520 520
 FT VARIANT 527 527
 FT SEQUENCE 746 AA; 83146 MW; EP913C5B4A0C8B5D8 CRC64;
 SQ
 Query Match 8.0%; Score 101.5; DB 1; Length 746;
 Best Local Similarity 22.1%; Pred. No. 0.92;
 Matches 50; Conservative 33; Mismatches 68; Indels 75; Gaps 10;
 27 YRRFVINGKLTSLAAGIGWPIGIAFGLLVIVGYGYEMWMAPIFSGIIFAGSTEMVIAL 86

CC STRAIN-1AM-M-101;
 RA Takahashi Y., Shonai F., Fujita Y., Kohchi T., Ohnaka K.,
 RA Matubara H.;
 RT "Structure of a co-transcribed gene cluster, *ndh1-ndh6-ndh4L*,
 RT cloned from the filamentous cyanobacterium *Plectononema boryanum*,";
 RL Plant Cell Physiol. 32:969-981(1991).
 CC -1- FUNCTION: Ndh-1 shuttles electrons from NAD(P)H, via FMN and iron-
 CC sulfur (Fe-S) centers, to quinones in the respiratory chain. The
 CC immediate electron acceptor for the enzyme in this species is
 CC believed to be plastoquinone. Couples the redox reaction to proton
 CC translocation (for every two electrons transferred, four hydroge
 CC ions are translocated across the cytoplasmic membrane), and thus
 CC conserves the redox energy in a proton gradient (By similarity).
 CC -1- CATALYTIC ACTIVITY: NAD(P)H + plastoquinone = NAD(P) (+) +
 CC plastoquinol.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -1- SIMILARITY: Belongs to the complex I subunit 1 family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: D01014; BAB00814.1;
 DR PIR: JQ2135; JQ2135.
 DR InterPro: IPR001694; Resp_NADH_dh1.
 DR Pfam: PF00146; NADHdh_1.
 DR PROSITE: PS00667; COMPLEX1_ND1_1; FALSE_NEG.
 DR PROSITE: PS00668; COMPLEX1_ND1_2; 1.
 KW Oxidoreductase; NAD; NADP; Quinone; Plastoquinone; Transmembrane.
 FT TRANSMEM 29 51 POTENTIAL.
 FT TRANSMEM 97 119 POTENTIAL.
 FT TRANSMEM 129 151 POTENTIAL.
 FT TRANSMEM 254 276 POTENTIAL.
 FT TRANSMEM 305 327 POTENTIAL.
 FT TRANSMEM 348 370 POTENTIAL.
 SQ SEQUENCE 372 AA; 40239 MW; 974729ED98F2094D CRC64;
 Query Match 8.7%; Score 110.5; DB 1; Length 372;
 Best Local Similarity 23.2%; Pred. No. 0.096; Mismatches 80; Indels 99; Gaps 17;
 Matches 66; Conservative 39; Mismatches 80; Indels 99; Gaps 17;
 QY 12 EVSPKALAPDDKGRYREIAQGLTSLAAGLGMVPIGALVIOGYEWMAAPFS 71
 DB 88 DVVPAKA-----DK-----LLFTLGPALVIVPFLSY--LILPQGQLQITDVL 130
 QY 72 GLIFAGSTMLVIALVVGAPLGAIA-----LTLVNFHRYVYAFSPFLVYKAEIA 124
 DB 131 GI-----FLWIAL--SSVPIGLIMSGYASNNKXSLGLGRMAAQSSYEL-----PLA 177
 QY 125 RFYSVPAL-----DEAVYATAPAGMSAMRLISMQIAFHSYVFGGLTG 170
 DB 178 --LSVLAIVMGNSLSTVDIVNQAQYGI-----LGMNWR--QPVGIIIFWI-----AA 223
 QY 171 VALAEIIPF-----EIKLEFALCSI--FVTLTLDSCRKQIPSLIAG 213
 DB 224 LAECERIPFDLPABEBELVAGVOTESGMKFPALFYGSYVNTLSAL-----LFAVLVIGG 279
 QY 214 LSEFTALVVIPIG-----QALFALLIFLGLLTIRYFPL 246
 DB 280 WEPIPSLSVIGLIGVPESTPWLQILFAT--IGIGMTLLKAYFL 321

DE NAD(P)H-quinone oxidoreductase chain 4, chloroplast (EC 1.6.5.-)
 DE (NAD(P)H dehydrogenase, chain 4) (NADH-plastoquinone oxidoreductase
 DE chain 4).
 GN NDH4.
 OS Psilotum nudum (Whisk fern).
 OG Chloroplast.
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Moniliformopses; Psilotophyta; Psilotales; Psilotaceae; Psilotum.
 CC NCBI TaxID=3240;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Kinyokyu;
 RA Waksugi T., Nishikawa A., Yamada K., Sugiura M.;
 RT "Complete nucleotide sequence of the chloroplast genome from a fern,
 RT *Psilotum nudum*,";
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: NAD(P)H + plastoquinone = NAD(P) (+) +
 CC plastoquinol.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AP004638; BAB84274.1; ALT_INIT.
 DR HAMAP: MF_00491; 1.
 DR InterPro: IPR003918; NADHdh_oxred4.
 DR InterPro: IPR001750; Oxidored_g1.
 DR Pfam: PF00361; oxidored_g1; 1.
 DR PRINTS: PR01437; NDOXDRDASE4.
 KW Oxidoreductase; NAD; NADP; Quinone; Plastoquinone; Chloroplast.
 SQ SEQUENCE 498 AA; 55958 MW; 7C89E6003B8EAD95 CRC64;
 Query Match 8.0%; Score 101.5; DB 1; Length 498;
 Best Local Similarity 22.5%; Pred. No. 0.64; Mismatches 81; Indels 117; Gaps 16;
 Matches 69; Conservative 40; Mismatches 81; Indels 117; Gaps 16;
 QY 44 LGWY--PIGIAFGILVIOGYEW-----WAP--LFGSLIFAGSTEMVIAL 86
 DB 119 LAMVSGIGLFAAGNILLFFPMWELIPYLLSMWGGKRLVSSKFLYTAGSIFL 176
 QY 87 VVGAPLG-----YVFLIDBAVAVTAPAGMSAMRLISMQIAFHSYVFG-----GL 168
 DB 179 LVGSLTGWGLVGSNGTSFDLITNRSYPIVETALY-----FSFLIYVAVKLPILP 230
 QY 127 -----YSVFLIDBAVAVTAPAGMSAMRLISMQIAFHSYVFG-----GL 168
 DB 231 HTWLPDTHGHAHISTCML-----AGILLKGGVGLR--INMELLPHAHF--FSPFLVYMG 285
 QY 169 TGAVALIEIPFIEKLE--FALCSL-----FVTLTLDSCR--TKQIPSL-- 210
 DB 286 VQIYASLISFSLNIRKRIAYSSHMGVVIIGISITDIGNGAILQWISHGLIGAL 345
 QY 211 --LAGLS-----FTIALVVP--PGQALFAA--LLIFLGL 239
 DB 346 FFLAGISYDRTQTLVLDQGLIAPVMPRLPTMFSFSLASLALPGSSFAEFLIFLGI 405
 QY 240 TIRYFPL 246
 DB 406 TNPYSL 412

RESULT 9

NCUC_HUMAN
 ID CLC5_HUMAN STANDARD; PRT; 746 AA.
 AC P51795;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DS Chloride channel protein 5 (CLC-5).

SEQUENCE FROM N.A.
 RC STRAIN=K12 / MG1655;
 RA MEDLINE=97426617; PubMed=9278503;
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirtpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Shao Y.;
 RA "The complete genome sequence of *Escherichia coli* K-12";
 RL Science 277:1453-1474(1997).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -1- SIMILARITY: Belongs to the azic family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AE000353; AAC75729.1; -
 CC PIR: C65048; C65048.
 CC EcoGene: EG13528; YgaZ.
 DR InterPro: IPR004471; Azic.
 DR Pfam: PF03591; Azic; 1.
 KW Hypothetical protein; Transport; Transmembrane; Complete proteome.
 FT TRANSMEM 25 45 POTENTIAL.
 FT TRANSMEM 64 84 POTENTIAL.
 FT TRANSMEM 110 130 POTENTIAL.
 FT TRANSMEM 141 161 POTENTIAL.
 FT TRANSMEM 173 193 POTENTIAL.
 FT TRANSMEM 206 226 POTENTIAL.
 FT SEQUENCE 245 AA; 26107 MW; 22ACE8A8BD7D651B CRC64;
 SQ
 Query Match 15.5%; Score 197; DB 1; Length 245;
 Best Local Similarity 29.1%; Pred. No. 1,1e-08; Indels 30; Gaps 9;
 Matches 68; Conservative 41; Mismatches 95;
 QY 34 QGLKTLAAGLGMVPIGIAFGILVIOGYEWMAAPFSGILFAGSTEMLVIALVGAAPL 93
 DB 17 EGCKDLPIVISTYIPAFAGLGNATRLGSPLESVFSCIIYAGASQFVTMLAAGSL 76
 QY 94 GAIALTLTLVNFPHVYAFSPFLAV- --KPIAFYVFPALIDBAVNTAR- --PA 145
 DB 77 WIALVTVMADVHVHYGSLRSRIIQLQSKRTALW- --AFGLTEVEFPAATAKVRNR 134
 QY 146 GMSAMRLISMQIAFHSYVW- --GGILGVAIAELIPFIKIGLEFALCSLFTLLTDSR 201
 DB 135 RSEBNMMIGIAFSMSWSWGTGVIAGFSGSLQGVPAVBALGFLPALFMSFLLASFO 194
 QY 202 TKKQISLLAAGLSFTIALV-VIPGQALPA- --ALI- --IFLGLT- --IYFPLG 247
 DB 195 RKQ- --SLCVTAALVGAAGVTLFSIPVAIIAGIVCGCLTALIOAFWQG 240
 RESULT 6
 YH55_ARCFU STANDARD; PRT; 219 AA.
 ID YH55_ARCFU
 AC 028519;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical protein AF1755.
 GN AF1755.
 OS Archaeoglobus fulgidus.
 OC Archaea; Euryarchaeota; Archaeoglobi; Archaeoglobales;
 OC Archaeoglobaceae; Archaeoglobus.
 ON NCBI_Taxid=2234;
 RX STRAIN=VC-16 / DSM 4304 / ATCC 49558;
 RX MEDLINE=98049343; PubMed=9389475;
 RA Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,

RA Ketchum K.A., Dodson R.J., Gwin M., Hickey E.K., Peterson J.D.,
 RA Richardson D.L., Kierlavage A.R., Graham D.E., Kyriides N.C.,
 RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
 RA Kirschner E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
 RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodex A., Zhou L.,
 RA Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Uitterback T.,
 RA Cotton M.D., Spriggs T., Arlrich P., Kaine B.P., Sykes S.M.,
 RA Sadow P.M., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
 RA Mason T.M., Olsen G.T., Fraser C.M., Smith H.O., Woese C.R.,
 RA Venter J.C.;
 RA "The complete genome sequence of the hyperthermophilic, sulphate-
 RT reducing archaeon *Archaeoglobus fulgidus*.";
 RL Nature 390:364-370(1997).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -1- SIMILARITY: Belongs to the azic family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AE000982; AAB89495.1; -
 CC PIR: B69469; B69469.
 DR TIGR: AF1755; -
 DR InterPro: IPR004471; Azic.
 DR Pfam: PF03591; Azic; 1.
 KW Hypothetical protein; Transport; Transmembrane; Complete proteome.
 FT TRANSMEM 14 34 POTENTIAL.
 FT TRANSMEM 37 57 POTENTIAL.
 FT TRANSMEM 123 143 POTENTIAL.
 FT TRANSMEM 155 175 POTENTIAL.
 FT TRANSMEM 189 209 POTENTIAL.
 FT SEQUENCE 219 AA; 23337 MW; 265A2C88DA93EBB6 CRC64;
 SQ
 Query Match 11.4%; Score 145; DB 1; Length 219;
 Best Local Similarity 24.2%; Pred. No. 0.00012; Indels 36; Gaps 8;
 Matches 56; Conservative 43; Mismatches 96;
 QY 14 SPSKALEPDDKGRVREIAQGLKTSIAAGLGMVPIGIAFGILVIOGYEWMAAPFSG 73
 DB 3 SLKSGAM- --FKGLVYSPPTVMATYIPAFVFTGVALRTLGFEVBNMLASL 51
 QY 74 IFAGSTEMLVIALVGAAPGAIATLTLVNFHRYA- --FSPFLHYVKKPIAR 125
 DB 52 IFAGASOFALITL-YQSGLISAIFI-PIPLNHRHIIYSSIIAOKLKLRF-HI- -- 101
 QY 126 FYSVFALIDEAVYATARPPAGMSAMRLISMQIAFHSYVWGGTLGVAIALIPE- --IK 182
 DB 102 --SAFGLTDEVFAVS- --NSAENRFLGLGELSGYSANWGTLGVALGTLIDRIVYS 157
 QY 183 GLERFALCSLFTLLTDSCKTKQIPSLLAGLSFTIALV-VIPGQALPAAL 233
 DB 158 ALVFESISALFVLLPLNKGHRVRAAVSGAVLAFHLLMLTSGIILAAAL 208
 RESULT 7
 N01C_PLEBO STANDARD; PRT; 372 AA.
 ID N01C_PLEBO
 AC Q00242;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE NAD(P)H-quinone oxidoreductase chain 1 (EC 1.6.5.-) (NAD(P)H
 DE dehydrogenase 1, chain 1) (NDH-1, chain 1).
 GN NDHA OR NDH1.
 OS Plectonema boryanum.
 OC Bacteria; Cyanobacteria; Oscillatoriales; Plectonema.
 ON NCBI_Taxid=1184;
 RX N
 RX SEQUENCE FROM N.A.

RN SEQUENCE FROM N.A.
 RX MEDLINE=99120557; PubMed=9923682;
 RA Alm R.A., Ling U.-S.L., Noir D.T., King B.L., Brown E.D., Doig P.C.,
 RA Smith D.R., Noonan B., Guild B.C., deJonge B.L., Carmel G.,
 RA Tummino P.J., Carnuo A., Urie-Nickelsen M., Mills D.M., Ives C.,
 RA Gibson R., Merberg D., Mills S.D., Jiang Q., Taylor D.E., Vovis G.F.,
 RA Trust T.J.;
 RT "Genomic sequence comparison of two unrelated isolates of the human
 RT gastric pathogen *Helicobacter pylori*."
 RL Nature 397:176-180(1999).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -1- SIMILARITY: Belongs to the azic family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AE001548; AAD06824.1; -
 CC PIR; F71831; F71831.
 DR InterPro; IPR004471; Azic.
 DR Pfam; PF03591; Azic; 1.
 DR TIGRFAMs; TIGR00346; azic; 1.
 KM Hypothetical protein; Transport; Transmembrane; Complete proteome.
 FT TRANSMEM 14 34 POTENTIAL.
 FT TRANSMEM 42 62 POTENTIAL.
 FT TRANSMEM 130 150 POTENTIAL.
 FT TRANSMEM 156 176 POTENTIAL.
 FT TRANSMEM 192 212 POTENTIAL.
 SQ SEQUENCE 228 AA; 25844 MW; 5D6EBA86BE2B9744 CRC64;
 Query Match 18.1%; Score 229.5; DB 1; Length 228;
 Best Local Similarity 29.6%; Pred. No. 2.8e-11;
 Matches 61; Conservative 41; Mismatches 95; Indels 9; Gaps 5;
 QY 44 LGMYDGIAGFLVLYQYGEWMAAPLPSGLIFAGSTEMLYALVGAAPLGAIALTLTLV 103
 DB 20 LGYLLMGMTFGWMLAQOGDYKVALFMSLFIYAGIQVATLLSAQASLNNVIVSLV 79
 QY 104 NRRHVFYAFSPFLHVKNPPIARF-VSPFALIDEAVAVT---AARPAQSAWRLISMQIAF 159
 DB 80 NARQCIALSW-LDFPKTKTKRPLFLAALHDETALNLNYPKGVNETPMSISILN 138
 QY 160 HSYWVFGLTGVVAIAELIPEIKGLEFALCSLFTVLLDSCRTKQIPSLLAGISFTIA 219
 DB 139 HSYWVFGLTGVVAIAELIPEIKGLEFALCSLFTVLLDSCRTKQIPSLLAGISFTIA 219
 QY 220 LVVIRPGQALF---ALLIFLGLTLTR 242
 DB 198 CLAPFTEYFLLIALVLMVLAFLIFR 223
 RESULT 4
 YH38_HAEIN STANDARD; PRT; 244 AA.
 ID_YH38_HAEIN
 AC P44302;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Hypothetical protein H11738.
 GN H11738.
 OS Haemophilus influenzae.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
 OC Pasteurellaceae; Haemophilus.
 OX NCB1_TaxID=727;
 OX [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Rd / KM20 / ATCC 51907;
 RX MEDLINE=95350630; PubMed=7542800;

RA Fleischmann R.D., Adams K.D., White O., Clayton R.A., Kirkness E.F.,
 RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
 RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
 RA Scott J.D., Shiley R., Liu L.-I., Glodek A., Kelley J.M.,
 RA Weidman J.F., Phillips R., Spillig T., Hedblom E., Cotton M.D.,
 RA Urtreback T.R., Hanna M.C., Spriggs T., Saudek D.M., Brandon R.C.,
 RA Fine L.D., Fritchman J.D., Fritchman J.L., Geoghagen N.S.M.,
 RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
 RA Venter J.C.;
 RT "Whole-genome random sequencing and assembly of *Haemophilus influenzae*
 RT Rd".
 RL Science 269:496-512(1995).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -1- SIMILARITY: Belongs to the azic family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; U32846; AAC23382.1; -
 CC PIR; P64041; P64041.
 DR TIGR; H11738; -
 DR InterPro; IPR004471; Azic.
 DR Pfam; PF03591; Azic; 1.
 DR TIGRFAMs; TIGR00346; azic; 1.
 KM Hypothetical protein; Transport; Transmembrane; Complete proteome.
 FT TRANSMEM 21 41 POTENTIAL.
 FT TRANSMEM 44 64 POTENTIAL.
 FT TRANSMEM 66 86 POTENTIAL.
 FT TRANSMEM 139 159 POTENTIAL.
 FT TRANSMEM 165 185 POTENTIAL.
 FT TRANSMEM 199 219 POTENTIAL.
 SQ SEQUENCE 244 AA; 27360 MW; 70C61B72D4AF16021 CRC64;
 Query Match 17.9%; Score 227.5; DB 1; Length 244;
 Best Local Similarity 27.5%; Pred. No. 4.3e-11;
 Matches 57; Conservative 50; Mismatches 87; Indels 13; Gaps 6;
 QY 45 GMYPIGIAGFLVLYQYGEWMAAPLPSGLIFAGSTEMLYALVGAAPLGAIALTLTLV 103
 DB 30 GFLFLGIAGIYKALGFGFLYPTLMALIVGVSFFIAGALIAFPSPV-SVLLITLMT 88
 QY 104 NRRHVFYAFSPFLHVKNPPIARF-VSPFALIDEAVAVT---VTARPAQSAWRLISMQ 156
 DB 89 SANQIFRIGISMLEKIGIHGK---RWYLLITLVDSFSLNWKAKIPPHLDKGMVFFVS 145
 QY 157 IAPHSYWVFGLTGVVAIAELIPEIKGLEFALCSLFTVLLDSCRTKQIPSLLAGISFTIA 215
 DB 146 IYHIVWVVGAAAGNLFGLVLPNLKGVFSMTALFLVFAENWMLKQSHSSLLGLGIA 205
 QY 216 FTIALVIRPGQALFALLIFLGLTLTR 242
 DB 206 LVFLILIGKXYFLIPTIGIWLITMK 232
 RESULT 5
 YGAZ_ECOLI STANDARD; PRT; 245 AA.
 ID_YGAZ_ECOLI
 AC P76630;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical protein YGAZ.
 GN YGAZ OR B2682.
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.
 OX NCB1_TaxID=562;
 OX [1]

SEQUENCE FROM N.A.
 RC STRAIN=K12 / MG1655;
 RX MEDLINE=97426617; PubMed=9278503;
 RA Blatter F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-Verdes J., Glesner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.M., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Shao Y.,
 RT "The complete genome sequence of *Escherichia coli* K-12.",
 RL Science 277:1453-1474(1997).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -1- SIMILARITY: Belongs to the azic family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL, AB000353; AAC75729.1; -
 CC PIR, C65048; C65048.
 CC DR Ecocore; EG13528; Y98Z.
 CC DR InterPro; IPR004471; Azic.
 CC Pfam; PF03591; Azic; 1.
 CC KW Hypothetical protein; Transport; Transmembrane; Complete proteome.
 CC FT TRANSMEM 25 45 POTENTIAL.
 CC FT TRANSMEM 64 84 POTENTIAL.
 CC FT TRANSMEM 110 130 POTENTIAL.
 CC FT TRANSMEM 141 161 POTENTIAL.
 CC FT TRANSMEM 173 193 POTENTIAL.
 CC FT TRANSMEM 206 226 POTENTIAL.
 CC SEQUENCE 245 AA; 26107 MW; 22ACE8ABD7D651B CRC64;
 SO

Query Match 15.5%; Score 197; DB 1; Length 245;
 Best Local Similarity 29.1%; Pred. No. 1.1e-08; Mismatches 95; Indels 30; Gaps 9;
 Matches 68; Conservative 41; Mismatches 95; Indels 30; Gaps 9;

QY 34 OGLKTSLAAGLQMPYPIGAFGLVIOGYEWMAAPLFSGLIPAGSTEMLVIALVGAAPL 93
 DB 17 EGCKSLPIVIVISYIPVAFAGLNATRLGSPLESVFSCIIYAGASQVITMLAAGSSL 76
 QY 94 GAIALTLTLVNPFRHYVYASFPPLAV- - -KNPIARFYSVPLIDEAVALAAR- - -PA 145
 DB 77 WIAALTVKMDVRAHVLYGSLRSRIIQLQSKTLW- -AFGLTDEVRAATAKLVKRR 134
 QY 146 GMSAMRLISMOIAFHSYWF- - -GGLTGVAIAEELIPFRIKLEFALCSLPFTLLDSCR 201
 DB 135 RMSENMWIGIAFSSMSWFGVTVIGAFSGSGLLGYPAVEALGFMLPALFWSFLASFO 194
 QY 202 TKKQIPLSLLAGISFTIALV-VIPQALPA- - -ALL- -IFGLTLT- -IRYFELG 247
 DB 195 RKQ- - - -SLCVTALVGLAGVTLFSPVAILAGVCGCTALALQAEWQ 240

RESULT 6
 YHS5_ARCFU
 ID YHS5_ARCFU STANDARD; PRT; 219 AA.
 AC 028519;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical protein AF1755.
 GN AF1755.
 OS Archaeoglobus fulgidus.
 OC Archaea; Euryarchaeota; Archaeoglobi; Archaeoglobales;
 OC Archaeoglobaceae; Archaeoglobus.
 OX NCBI_TaxID=2234;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;
 RX MEDLINE=96049343; PubMed=9389475;
 RA Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,

Ketchum K.A., Dodson R.J., Gwinn M., Hickey E.K., Peterson J.D.,
 RA Richardson D.L., Kierlavage A.R., Graham D.E., Kyriades N.C.,
 RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
 RA Kiriess E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
 RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodet A., Zhou L.,
 RA Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Utterback T.,
 RA Cotton M.D., Spriggs T., Atliach P., Kaine B.P., Sykes S.M.,
 RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
 RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
 RA Venter J.C.,
 RT "The complete genome sequence of the hyperthermophilic, sulphate-
 RT reducing archaeon *Archaeoglobus fulgidus*.",
 RL Nature 390:364-370(1997).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -1- SIMILARITY: Belongs to the azic family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL, AB000982; AAB89495.1; -
 CC PIR, B69469; B69469.
 CC DR TIGR; AF1755; -
 CC DR InterPro; IPR004471; Azic.
 CC Pfam; PF03591; Azic; 1.
 CC KW Hypothetical protein; Transport; Transmembrane; Complete proteome.
 CC FT TRANSMEM 14 34 POTENTIAL.
 CC FT TRANSMEM 37 57 POTENTIAL.
 CC FT TRANSMEM 123 143 POTENTIAL.
 CC FT TRANSMEM 155 175 POTENTIAL.
 CC FT TRANSMEM 189 209 POTENTIAL.
 CC SEQUENCE 219 AA; 22337 MW; 265A2C88DA93EBB6 CRC64;
 SO

Query Match 11.4%; Score 145; DB 1; Length 219;
 Best Local Similarity 24.2%; Pred. No. 0.00012; Mismatches 96; Indels 36; Gaps 8;
 Matches 56; Conservative 43; Mismatches 96; Indels 36; Gaps 8;

QY 14 SPKRALPEPDDKGRYREINQGLKTSLAAGLQMPYPIGAFGLVIOGYEWMAAPLFSGL 73
 DB 3 SLKSNW- - - -FRKGLVSPFIYMAIIPVAFGLVARTLGSEVENAMLASL 51
 QY 74 IFAGSTEMLVIALVGAAPGATLTLVNPFRHYV- - - -FSPPLHYVKKPIAR 125
 DB 52 IFAGASQFALITL-YQSLSLALFI-PIPLNLSHIYSSIIAOKLRRP-HI- - - - 101
 QY 126 FYSVFALIDEAYAVTAARPAWGSAMRLISMOIAFHSYWFVGLTGVAIAELIPE- - -IK 182
 DB 102 - - -SAFGLTDEVFAVS- - -NSAENRFLGLGLSISYSAWVGTLGVLASTLLIDRVS 157
 QY 183 GLEFALCSLPVTLTLDSCRTRKQIPSLLAGISFTIALV-VIPQALPAAL 233
 DB 158 ALVFSISALFLVLLPVLKGRHVRAVSGAVALAFHLLNLTSGVIGIAPAL 208

RESULT 7
 NTIC_PLEBO
 ID NTIC_PLEBO STANDARD; PRT; 372 AA.
 AC Q00242;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE NAD(P)H-quinone oxidoreductase chain 1 (EC 1.6.5.-) (NAD(P)H
 DE dehydrogenase I, chain 1) (NDH-1, chain 1).
 GN NDH4 OR NDH1.
 OS Plectonema boryanum.
 OC Bacteria; Cyanobacteria; Oscillatoriales; Plectonema.
 OX NCBI_TaxID=1184;
 RN [1]
 RP SEQUENCE FROM N.A.

RN SEQUENCE FROM N.A.
 RP MEDLINE=99120557; PubMed=9923682;
 RX Alm R.A., Ling L.-S.L., Moir D.T., King B.L., Brown E.D., Doig P.C.,
 RA Smith D.R., Noonan B., Guild B.C., deGange B.L., Carmel G.,
 RA Tummino P.J., Caruso A., Uria-Nickelsen M., Mills D.W., Ives C.,
 RA Gibson R., Merberg D., Mills S.D., Jiang Q., Taylor D.E., Vovis G.F.,
 RA Trust T.J.;
 RT "Genomic sequence comparison of two unrelated isolates of the human
 RT gastric pathogen *Helicobacter pylori*."; Nature 357:176-180(1999).
 RL Nature 357:176-180(1999).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -1- SIMILARITY: Belongs to the azlC family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AB001548; AAC06824.1; -
 CC PIR: P71831; P71831.
 DR InterPro: IPR004471; AzlC.
 DR Pfam: PF03591; AzlC; 1.
 DR TIGRFAMs: TIGR00346; azlC; 1.
 KW Hypothetical protein; Transport; Transmembrane; Complete proteome.
 FT TRANSMEM 14 34 POTENTIAL.
 FT TRANSMEM 42 62 POTENTIAL.
 FT TRANSMEM 130 150 POTENTIAL.
 FT TRANSMEM 156 176 POTENTIAL.
 FT TRANSMEM 192 212 POTENTIAL.
 FT TRANSMEM 228 AA; 25844 MW; 5058485829744 CRC64;
 SQ SEQUENCE
 Query Match 18.1%; Score 229.5; DB 1; Length 228;
 Best Local Similarity 29.6%; Pred. No. 2.9e-11;
 Matches 61; Conservative 41; Mismatches 95; Indels 9; Gaps 5;
 QY 44 LGMTPIGAFGLVYQYGYEMWAPLFGSLFACSTEMVYALVGAAPLGAIALTLTV 103
 DB 20 LGYLTMGTFQMLAQQGYDYKALFMSLFTYAGAIQFVALITLLSAQASIMNVIVSLV 79
 QY 104 NERHVFYAFSPPLHVKPIARF-YVFALIDEAVAT---AAPAGWSARLISQIAF 159
 DB 80 NARQCYALSM-LDRFKTKRRLPYLAHLTDFETALLNLVAPKKGYNEDPMFSISLN 138
 QY 160 HSYWVFGGLTGAVALIPEIKGLEFALCSLFTVTLTDSCKTKQIPLSLAGLSTIA 219
 DB 139 HSYWVFGSLVSLVSGSHFSFDQGMFEVMTAIFVLFWMEQYKRNTHKNMWL-GIALAVV 197
 QY 220 LVVIFGQALF---ALLIFLGLTIR 242
 DB 198 CLALFGTEYFLILALVNLVALLIFR 223
 RESULT 4
 YH38_HABIN
 ID YH38_HABIN STANDARD; PRT; 244 AA.
 AC P44302;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Hypothetical protein H11738.
 GN H11738.
 OS Haemophilus influenzae;
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
 OC Pasteurellaceae; Haemophilus.
 OX NCB1_TaxID=727;
 RN SEQUENCE FROM N.A.
 RP STRAIN=rd / KW20 / ATCC 51907;
 RX MEDLINE=9550630; PubMed=7542800;

RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
 RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
 RA McKenny D., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
 RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
 RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
 RA Uiterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
 RA Pine L.D., Fritchman J.D., Fritchman J.L., Geoghagen N.S.M.,
 RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
 RA Venter J.C.;
 RT "Whole-genome random sequencing and assembly of *Haemophilus influenzae*
 RT Rd".
 RT Science 269:496-512(1995).
 RL Science 269:496-512(1995).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -1- SIMILARITY: Belongs to the azlC family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: U32846; AAC23382.1; -
 CC PIR: P64041; P64041.
 DR TIGR: H11738.
 DR InterPro: IPR004471; AzlC.
 DR Pfam: PF03591; AzlC; 1.
 DR TIGRFAMs: TIGR00346; azlC; 1.
 KW Hypothetical protein; Transport; Transmembrane; Complete proteome.
 FT TRANSMEM 21 41 POTENTIAL.
 FT TRANSMEM 44 64 POTENTIAL.
 FT TRANSMEM 66 86 POTENTIAL.
 FT TRANSMEM 129 159 POTENTIAL.
 FT TRANSMEM 165 185 POTENTIAL.
 FT TRANSMEM 199 219 POTENTIAL.
 FT TRANSMEM 244 AA; 27360 MW; 706187224AF16021 CRC64;
 SQ SEQUENCE
 Query Match 17.9%; Score 227.5; DB 1; Length 244;
 Best Local Similarity 27.5%; Pred. No. 4.3e-11;
 Matches 57; Conservative 50; Mismatches 87; Indels 13; Gaps 6;
 QY 45 GMYPIGAFGLVYQYGYEMWAPLFGSLFACSTEMVYALVGAAPLGAIALTLTV 103
 DB 30 GELFLGIAIGIYKALFGFLYPTLMALIVAGSEFPAAGLILAFSPF-SVLLITLMI 88
 QY 104 NERHVFYAFSPPLHVKPIARF-YVFALIDEAVAT---VTAPAGWGAARLISQ 156
 DB 89 SARQIVYGISLMKEKGIHIGKK---RWYITLVDSFESANTMAKIPHLDGWMTFVS 145
 QY 157 IAFHSYWVFGGLTGAVALIPEIKGLEFALCSLFTVTLTDS-CKTKQIPLSLAGIS 215
 DB 146 LVHITWVLAGAMGNLFTVLPFNKGVFMTALFLVFAFNWMLKGSHESSLGLGIA 205
 QY 216 FTALVIVIPQALFALLIFLGLTIR 242
 DB 206 LVFLITIGKXYFLIPTLIGIWLITWR 232
 RESULT 5
 YGAZ_ECOLI
 ID YGAZ_ECOLI STANDARD; PRT; 245 AA.
 AC P76630;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical protein YGAZ.
 GN YGAZ OR B2682.
 OS *Escherichia coli*.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; *Escherichia*.
 OX NCB1_TaxID=562;
 RN

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 16, 2004, 06:51:16 ; Search time 13.9833 Seconds

(without alignments)
934,659 Million cell updates/sec

Title: US-10-608-504-3

Perfect score: 1269

Sequence: 1 MCKTQIHSHLSLEVSFSPKAL.....LIIFGLLTIRYFLGKAK 251

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database: SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	240	18.9	254	1	AZLC_BACSU
2	235	18.5	228	1	YD31_HELPY
3	229.5	18.1	228	1	YD31_HELPY
4	227.5	17.9	244	1	YH38_HABIN
5	197	15.5	245	1	YGAZ_ECOLI
6	145	11.4	219	1	YH55_ARCFU
7	110.5	8.7	372	1	NUIC_PLEBO
8	101.5	8.0	498	1	NUAC_PLENU
9	101.5	8.0	746	1	CLC5_HUMAN
10	100.5	7.9	746	1	CLC5_MOUSE
11	100.5	7.8	746	1	CLC5_RAT
12	98.5	7.8	643	1	S212_HUMAN
13	98	7.7	760	1	CLC4_HUMAN
14	96.5	7.6	416	1	RHCE_HUMAN
15	95	7.5	747	1	CLC4_MOUSE
16	94	7.4	567	1	PROI_TORSE
17	93	7.3	266	1	PTNC_ECOLI
18	92	7.2	747	1	CLC4_RAT
19	91.5	7.2	464	1	GALP_ECOLI
20	91.5	7.2	695	1	YD38_YEAST
21	91	7.2	624	1	COX1_BACPF
22	90.5	7.1	226	1	ATP6_RHTUN
23	90.5	7.1	267	1	UPK_CAMEL
24	90.5	7.1	351	1	Y876_METUA
25	90.5	7.1	393	1	SETB_ECOLI
26	90	7.1	220	1	YICG_HABIN
27	90	7.1	416	1	RHD_HUMAN
28	90	7.1	416	1	RHLD_GORGO
29	90	7.1	528	1	TH12_TRYBB
30	90	7.1	528	1	TH2A_TRYBB
31	89.5	7.1	358	1	EMA3_CAMEL
32	89.5	7.1	319	1	PSTC_ECOLI
33	89.5	7.1	381	1	CYB_MACRO

34	89	7.0	428	1	BCH2_RHOCA	P26171 rhodobacter
35	88.5	7.0	226	1	ATP6_PHOVI	O00521 rhodospirillum rubrum
36	88.5	7.0	517	1	COX1_PISOC	P25001 pisaster oc
37	88.5	7.0	760	1	CLC3_CAVPO	O91279 cavla porce
38	88.5	7.0	760	1	CLC3_MOUSE	P51791 mus musculu
39	88.5	7.0	760	1	CLC3_RABIT	O18894 oryctolagus
40	88.5	7.0	760	1	CLC3_RAT	P51792 ratu
41	88.5	7.0	762	1	CLC3_HUMAN	P51790 homo sapien
42	88	6.9	144	1	BDHC_PSEFE	O8ghn3 pseudomonas
43	88	6.9	297	1	BACH_HALS4	O93741 halobacteri
44	88	6.9	362	1	NUIC_MAIZE	P25706 zea mays (m
45	88	6.9	499	1	NU4C_MARPO	P06263 marchantia

ALIGNMENTS

RESULT 1
AZLC_BACSU
ID 007942: STANDARD; PRT; 254 AA.
AC 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Branched-chain amino acid transport protein azlc.
GN AZLC OR BSU26710.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168 / BGSC1A1;
RX MEDLINE=97453479; PubMed=9287000;
RA Beltsky B.R., Gustafsson M.C.U., Sonenshein A.L., von Wachenfeldt C.,
RT "An lrp-like gene of Bacillus subtilis involved in branched-chain
RT amino acid transport." J. Bacteriol. 179:5448-5457 (1997).
RL (2)
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=97453479; PubMed=9308178;
RA Sorokin A., Bolotin A., Purnelle B., Hilbert H., Lauber J.,
RA Duesterhoeft A., Ehrlich S.D., Purnelle B., Hilbert H., Lauber J.,
RT "Sequence of the Bacillus subtilis genome region in the vicinity of
RT the lrp operon reveals two new extracytoplasmic function RNA
RT polymerase sigma factors SigV and SigZ." J. Microbiol. 143:2939-2943 (1997).
RL [3]
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=98044033; PubMed=9384377;
RA Kunz F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA Azevedo V., Berteiro M.G., Bessieres P., Bolotin A., Borchert S.,
RA Borriess R., Boursier L., Brane A., Braun M., Brigelli S.C., Bron S.,
RA Broiliet S., Brunsch C.V., Caldwell B., Capuano V., Carter N.M.,
RA Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,
RA Denizot F., Devine K.M., Duesterhoeft A., Ehrlich S.D., Emerson P.T.,
RA Euteneier K.D., Errington J., Fabre C., Ferrari E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
RA Ghim S.Y., Glaser P., Goffeau A., Goldlight E.J., Grandi G.,
RA Giuseppe G., Guy B.J., Haga K., Haleth J., Harwood C.R., Henaut A.,
RA Hilbert H., Holtsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
RA Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
RA Kobayashi Y., Koester P., Koningsreth G., Krogh S., Kumano M.,
RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
RA Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,
RA Medina N., Meilieu R.P., Mizuno M., Mosati D., Nakai S., Noack M.,
RA Noone D., O'Reilly M., Ogawa K., Ogawa A., Oudega B., Park S.H.,
RA Parro V., Pohl T.M., Portetelle D., Porwollik G., Prescott A.M.,
RA Prescan E., Puig P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA Rieger M., Rivolta C., Rocha B., Roche B., Rose M., Sadate Y.,
RA Sato T., Seaman B., Schleich S., Schroeter R., Scofield F.,
RA Sekiguchi J., Sekowska A., Serror S.U., Serror F., Shin B.S., Solde B.,

```

CC CC STRAIN=26695 / ANCC 700392; Clayton R.A., Sutton G.G.,
CC CC MEDLINE=973994467; PubMed=9252185; Nelson K.,
CC CC Tombl J.-F., White O., Kerlavage A.R., Kleink H.-P., Gill S., Dougherty B.A.,
CC CC Fleischmann R.D., Ketchum K.A., Klenk H.-P., Peterson S.,
CC CC Nelson K., Quackenbush J., Zhou L., Kirness E.F., Peterson S.,
CC CC Loftus B., Richardson D., Dodson R., Khakh H.G., Glodek A.,
CC CC McKenney K., Fitzgerald L.M., Lee N., Adams M.D., Hickey E.K.,
CC CC Berg D.E., Gooyane J.D., Utterback T.R., Peterson J.D., Kelley J.M.,
CC CC Cotton M.D., Weidman J.M., Fujii C., Bowman C., Matthey L., Wallin E.,
CC CC Hayes W.S., Borodovsky M., Karp P.D., Smith H.O., Fraser C.M.,
CC CC Venter J.C.;
CC CC "The complete genome sequence of the gastric pathogen Helicobacter
CC CC pylori."
CC CC Nature 388:539-547(1997).
CC CC -I- SUBCELLULAR LOCATION: Integral membrane protein (potential).
CC CC -I- SIMILARITY: Belongs to the azlC family.
CC CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC CC or send an email to license@isb-sib.ch).
CC CC -----
CC CC EMBL; AE000634; AACD08372.1; -.
CC CC DR PIR; C64686; C64686.
CC CC DR TIGR; HP1331; -.
CC CC DR InterPro: IPR004471; AzlC.
CC CC Pfam: PF03591; AzlC; 1.
CC CC TRIGRAMS; TIGR00346; azlC; 1.
CC CC KW Hypothetical protein; Transport; Transmembrane; Complete proteome.
CC CC FT TRANSMEM 14 POTENTIAL.
CC CC FT TRANSMEM 42 62 POTENTIAL.
CC CC FT TRANSMEM 130 150 POTENTIAL.
CC CC FT TRANSMEM 156 176 POTENTIAL.
CC CC FT TRANSEM 192 212 POTENTIAL.
CC CC SQ SEQUENCE 228 AA; 25854 MW; DB8EABBE8376130 CRC64;
CC CC -----
Query Match 18.5%; Score 235; DB 1; Length 228;
Best Local Similarity 33.0%; Pred. No. 1,Je-11;
Matches 70; Conservative 36; Mismatches 92; Indels 14; Gaps 7
Oy 44 LCWYPIGIAFGILVQGYEWMAAPLPISGLIFAGSTEMIVALTVPAGPYGIALTTLLV 103
Db 20 LGVILMGMTFGMLTVQGQGYDKVALFMSLFYGAVGFAVITLLSAQASIMNVIVSLIV 79
Oy 104 NRRAHYFAYSPFLHVKPKINRF-YSPFALIDEXAA--TAAPDAQMSMRI-SMQIAF 159
Db 80 NNRQTCTYALSM-LDRFKTKMRLEPLAHALIDEFPALLNYAPEKGVSEKDIPFSLSLN 138
Oy 160 HSAYWFGGLTGVAVALIELPFRIKGLEFALSFLTTLTDSCRTKKQIPSLLAGLSFTTA 219
Db 139 HSWIFGSLVGSIVGSHSPDTQGMFPMNAIFVLFMEDQ---YKRTHKNMAMLGIVIA 195
Oy 220 LVVIRGOALFPA--LIIFLGILTIIRYFFLK 248
Db 196 VVCL--ALFGTEYFLIALVLAIVLMLFRK 224
Cc -----
RESULT 3
YDJ3_HELPJ STANDARD; PRt. 228 AA.
ID_YDJ3_HELPJ
AC Q9ZUQ3;
DT 16-OCT-2001 (Rel. 40; Created)
DT 16-OCT-2001 (Rel. 40; Last sequence update)
DT 16-OCT-2001 (Rel. 40; Last annotation update)
DE Hypothetical protein JHP1251.
GN JHP1251.
OS Helicobacter pylori j99 (Campylobacter pylori j99).
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
OC Helicobacteraceae; Helicobacter.
NCBI_taxid=85963;

```

Tue Apr 20 06:47:18 2004

us-10-608-504-3.rpr

Page 6

```
Qy 28 RREYIAQCKTSLAAGLGMYPICIAFGILVIOGYEMWAPLPSGLIFAGSTEMVLIALV 87
Db 11 RETDMLRGAAIGSSIAIGYIPASLTFFGLAKGTGLTLETIAMSLEFYAGASQYMALTLT 70
Qy 88 VGAAPLGAIALTTLLVNFPHVFAFSPPLHV-KNPIARFYSVFALIDEAVAVTAARPAG 146
Db 71 AIGTGIEIILTTFTVINIHLMSASIRAKMEDTHPVKRAITAFGITDEVFALVTSQDR 130
Qy 147 WSAWRLISMOIAFHSYWFGLTGVAIAELIPPEI-KGLEFALCSLFTTLTIDSCRTKKQ 205
Db 131 LTNGFVIGVAVIAYVSWVHSVAVGYVGSALPATLQOGWGVALYAMFIALLIPSVKQRS 190
Qy 206 IPSELLAGLSFTIALVVIPOALFAL-----LIFGLLTRYFPFLGKAK 251
Db 191 V--LTLAG--TVALL-----NGLFSLFPSPGWSIILATLILASVGEIGKAK 233
```

Search completed: April 16, 2004, 06:54:24
Job time : 21.5766 secs

Cy

144 PAFGWSAMRLISMQIAFHSHYWFVGGTLTGAIA-----ELIPFEKIGLGFALCSLFVLITD 198
 :: :
Db 135 DGRSENNWMLGIATVSMTLSWLGTAIAGMFGNPLENPDAIEASISFMLEPFLSFLLA 194

Oy

199 SCRTCKOIFBULLAGS-----FTIALVTVPQA 227
 ||| :
195 S--FKRQYSILVIASLGALLGVLPSTIPAIIAGIA 229

RESULT 14

E86826
amino acid permutase yqfJ [imported] - Lactococcus lactis subsp. lactie (strain IL1403)
C:Species: Lactococcus lactis subsp. lacties
C>Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 03-Aug-2001
C:Accession: E86826
R:Bolocin, A.; Munkner, P.; Mauger, S.; Jalllon, O.; Malmer, K.; Weissenbach, J.; Ehrlich,
Gene Res. 11, 731-753, 2001
A>Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis ss
I:Reference number: AB6625; MUID:21235186; PMID:11337471
A:Accession: E86826
A>Status: Preliminary
A:Molecule type: DNA
A:Residues: 1..235 <STO>
A:Cross-references: GB:E005176; PID:g12724620; PIDN:AAR05711.1; GSPDB:GN00146
A:Experimental source: strain IL1403
C:Genetics:

A:Gene: yqfJ

Query Match

Best Local Similarity 13.8%; Score 174.5; DB 2; Length 235;
Matches 61; Conservative 47; Mismatches 92; Indels 39; Gaps 9;

Cy

34 QGLTKSLAAGLMGPPIGATFGLLVIOGYEMWAAPFESGLFPGASTETMLVTALVVGAAPL 93
 :::|||||
Db 9 QGLKDTPPTFYTGITGTAFCMGHSGEFSFWIVILLSLTVAGSAQFIWMSLATHSPI 68
 :
Cy 94 GAIALTTLLVFPRHVFAFSFPPLHAVKNPIARFYSVALI-DEAYAV----- 139
 :::|||:|
Db 69 MSIVSVFLVNSRIIMSMTTASYFNESLIKKNIILTGLTDSEFFALGMNKONYTEGKLN 126
 :
Cy 140 ----TARPAGWSAMRLISMQIAFHSHYWVGGLGVAAIELPEIKGLEPALCSLFV-- 193
 :
Db 129 PSWNASNLAIVLWALS-----AIGALIIGNLIA-N-P-EKGLEGHVAIAMITGL 176
 :
Cy 194 -LTLLDSCTRKQOIFSULLAGS---FTIALVTVPQG-ALPAALT-IFTGLLTRYYF 245
 :
Db 177 LVQLSIDKTGLTMLQVMWGITLVLYFYGILFPSNLVLEVTLICACAGVKRVFF 235
 :

RESULT 15

F84013
branched-chain amino acid transporter BH2910 [imported] - Bacillus halodurans (strain C
C:Species: Bacillus halodurans
C>Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 01-Mar-2002
C:Accession: F84013
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Saasaki, R.; Masui, N.; Fujii, F.; Hiro
Nutricl Acids Res. 28, 4317-4331, 2000
A>Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A:Reference number: AB3650; MUID:20512582; PMID:11058132
A:Accession: F84013
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1..237 <STO>
A:Cross-references: GB:AD001517; GB:BA000004; NID:g10155500; PIDN:BB06629.1; GSPDB:GNU
A:Experimental source: strain C-125
C:Genetics:

A:Gene: BH2910
C:Superfamily: hypothetical protein b2682

Query Match

Best Local Similarity 13.8%; Score 174.5; DB 2; Length 237;
Matches 55; Conservative 48; Mismatches 112; Indels 17; Gaps 6;

QY 88 VCAAPLGAIALTLTLVNFRRHVFYAFSPPLHVKNPIA-----RFYSVFALIDEAY 137
 A:Gene: VCAL002
 A:Map position: 2
 C:Superfamily: hypothetical protein b2682

Db 66 THPAAWTLGFPALLVNRHVLMSAS-----IAGKLDDFKGMRKXKAMLVLTDESW 116

QY 138 AVTAAR-----PAGWSANRLISMOIAFHSYVVEGGLTGVAAIEIPEIIGLEFALCSL 191
 A:Gene: VCAL002
 A:Map position: 2
 C:Superfamily: hypothetical protein b2682

Db 117 ALAEFRVIAGRLTAFFAGALSIYLVNLTATLGAALLGAVVGDV---SVIGLDPAFPAY 173

QY 132 FVTLTIDSCRTKQIPSLILAGLSFTIALVVPQALPAA 231
 A:Gene: VCAL002
 A:Map position: 2
 C:Superfamily: hypothetical protein b2682

Db 174 FIVLMGFVKGRGTGVLLASASACLTHTALVPGAWYIAA 213

RESULT 9

B97492
 azic family protein [imported] - Agrobacterium tumefaciens (strain C58, Cereon)
 C:Species: Agrobacterium tumefaciens
 C:Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 18-Nov-2002
 C:Accession: B97492
 R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman, A.; Liu, F.; Wollam, C.; Allinger, M.; Doughy, D.; Scott, C.; Lappas, C.; Markelz, B.; Science 294, 2323-2328, 2001
 A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tumefaciens
 A:Reference number: A97359; PMID:21608551; PMID:11743194
 A:Accession: B97492
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-234 <KUR>
 A:Cross-references: GB:AE007869; PIDN:AAK86891.1; PID:g15156113; GSPDB:GN00169
 C:Gene: AGR C 2001
 A:Map position: circular chromosome
 C:Superfamily: hypothetical protein b2682

Query Match 14.3%; Score 181; DB 2; Length 234;
 Best Local Similarity 25.0%; Pred. No. 3.6e-08;
 Matches 55; Conservative 46; Mismatches 91; Indels 28; Gaps 4;

QY 28 RREYINAGLTSLAGIGWPIGIAGLVLYQGYEMWAAPIPSGILPAGSTEMVIALY 87
 A:Gene: VCAL002
 A:Map position: 2
 C:Superfamily: hypothetical protein b2682

Db 6 KRAELIAGLSAAPLVAWPVIGVFGAVALIGKLSPLKSLMLVLPAGSGQFVAMDLW 65

QY 88 VCAAPLGAIALTLTLVNFRRHVFYAFSPPLHVKNPIA-----RFYSVFALIDEAY 137
 A:Gene: VCAL002
 A:Map position: 2
 C:Superfamily: hypothetical protein b2682

Db 66 THPAAWTLGFPALLVNRHVLMSAS-----IAGKLDDFKGMRKXKAMLVLTDESW 116

QY 138 AVTAAR-----PAGWSANRLISMOIAFHSYVVEGGLTGVAAIEIPEIIGLEFALCSL 191
 A:Gene: VCAL002
 A:Map position: 2
 C:Superfamily: hypothetical protein b2682

Db 117 ALAEFRVIAGRLTAFFAGALSIYLVNLTATLGAALLGAVVGDV---SVIGLDPAFPAY 173

QY 192 FVTLTIDSCRTKQIPSLILAGLSFTIALVVPQALPAA 231
 A:Gene: VCAL002
 A:Map position: 2
 C:Superfamily: hypothetical protein b2682

Db 174 FIVLMGFVKGRGTGVLLASASACLTHTALVPGAWYIAA 213

RESULT 10

G82388
 azic family protein VCA1002 [imported] - Vibrio cholerae (strain N16961 serogroup O1)
 C:Species: Vibrio cholerae
 C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 01-Mar-2002
 C:Accession: G82388
 R:Heideberg, U.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gilm, M.L.; Dodson, R.J.; Chardson, D.; Ermolaeva, W.D.; Vamathevan, J.; Bae, S.; Qin, H.; Dragoi, I.; Sellers, J.; R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
 A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
 A:Reference number: A82035; PMID:20406833; PMID:10952301
 A:Accession: G82388
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-239 <HEI>
 A:Cross-references: GB:AE004427; GB:AE003853; NID:G9658442; PIDN:AAF96898.1; GSPDB:GN001

A:Experimental source: serogroup O1; strain N16961; biotype El Tor
 C:Gene: VCAL002
 A:Map position: 2
 C:Superfamily: hypothetical protein b2682

Query Match 14.3%; Score 181; DB 2; Length 239;
 Best Local Similarity 23.6%; Pred. No. 3.7e-08;
 Matches 48; Conservative 47; Mismatches 104; Indels 4; Gaps 4;

QY 34 QGKTSIAGLGWPIGIAGLVLYQGYEMWAAPIPSGILPAGSTEMVIALYVGAAL 93
 A:Gene: VCAL002
 A:Map position: 2
 C:Superfamily: hypothetical protein b2682

Db 20 QGTAMLPISIALVPMGLLAGSFAIEAGLSVIESQALSAYLAGAQLVAGIEMFTGAGL 79

QY 94 GATLTLVNFRRHVFYAFSPPLHVKNPIA-----RFYSVFALIDEAYAVTAAR--PAGWSANR 151
 A:Gene: VCAL002
 A:Map position: 2
 C:Superfamily: hypothetical protein b2682

Db 80 LSLIATPFTSHRFLYSVMSRKSIEPLPLRMILTLGFLITDELPAICAGSDKQFNRY 139

QY 152 LISMOIAFHSYVVEGGLTGVAAIEIPEIIGLEFALCSLFTVTLTIDSCRTKQIPSL 210
 A:Gene: VCAL002
 A:Map position: 2
 C:Superfamily: hypothetical protein b2682

Db 140 ALGAGLSFYLIWMILASLVAGSYLPDLNQMGLPEPAVATFPAIVIPNIKSWPLISVL 199

QY 211 LA-GLSTIALVVPQALPAA 232
 A:Gene: VCAL002
 A:Map position: 2
 C:Superfamily: hypothetical protein b2682

Db 200 TALVLSVLTVMGIEGSLMFAST 222

RESULT 11

G95016
 conserved hypothetical protein SP0146 [imported] - Streptococcus pneumoniae (strain TIGR
 C:Species: Streptococcus pneumoniae
 C:Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 03-Aug-2001
 C:Accession: G95016
 R:Johnson, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heik
 on, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapple,
 neon, T.; Hickey, E.K.; Holt, I.E.
 A:Title: Complete Genome Sequence of a Virulent Isolate of Streptococcus pneumoniae.
 A:Reference number: A95000; PMID:21357209; PMID:11463916
 A:Accession: G95016
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-218 <KUR>
 A:Cross-references: GB:AE00672; PIDN:AAK74328.1; PID:g14971611; GSPDB:GN00164; TIGR:SP4
 A:Experimental source: strain TIGR4
 C:Gene: SP0146

Query Match 14.0%; Score 177.5; DB 2; Length 218;
 Best Local Similarity 26.6%; Pred. No. 6.7e-08;
 Matches 63; Conservative 36; Mismatches 82; Indels 39; Gaps 11;

QY 40 LAAGLWPIGIAGLVLYQGYEMWAAPIPSGILPAGSTEMVIALYVGAALP 95
 A:Gene: VCAL002
 A:Map position: 2
 C:Superfamily: hypothetical protein b2682

Db 1 METALGVYSIGAGGIIAGPY-----VTPVEMGLMSLFVYAGSAQFALALIVQAPVAA 55

QY 96 IALTTLVNFRRHVFYAFSPPLHVKNPIA-----PALIDEAYAVTAARAG----- 146
 A:Gene: VCAL002
 A:Map position: 2
 C:Superfamily: hypothetical protein b2682

Db 56 IAMTFLINLR-----LFLSLHASTYFHTSLMWNIGMSILTBETTVLVNGELAHDKV 111

QY 147 ---WSANRLISMOIAFHSYVVEGGLTGVAAIEIPEIIGLEFALCSLFTVTLTIDSCR- 201
 A:Gene: VCAL002
 A:Map position: 2
 C:Superfamily: hypothetical protein b2682

Db 112 NPMWMMGNLNSXYA-----WFGVGVGALGLLPNPEIFGLDPALVGMFTGIFASOPQI 167

QY 202 TKQIP-----SLLAGLSFTIALVVPQALPAA 233
 A:Gene: VCAL002
 A:Map position: 2
 C:Superfamily: hypothetical protein b2682

Db 168 MORRIPVNNLITLIVAVVSFPLTVMWSGLAVLFAILL 207

RESULT 12

H83390
 hypothetical protein PA2039 [imported] - Pseudomonas aeruginosa (strain PA01)

Db 101 NARHLYGISMIDKRYGTGKKLYLIFGWCDESFISINTANVPANVDKGMFVTLTNN 160
 Qy 161 SYWVWGGLTGVAAIAELIPEIKGLEFALCSLFTVTLTDSCTRKQIPSLLAGSFTIA 220
 Db 161 LTVWVGAALIGIFGSGVYKENTGDLDPVMTALFIVLFIEQMKKEKHYS-ALTGLGLSYAS 219
 Qy 221 VVIPG--QALFALLIFGLLTI 241
 Db 220 LILFGNGQFTIPAMLAIGVLT 242

RESULT 2

Conserved hypothetical integral membrane protein HP1331 - Helicobacter pylori (strain 26 C64686)
 C/Species: Helicobacter pylori
 C/Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jun-2000
 C/Accession: C64686
 R/Tom: J.F.; White, O.; Keriavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.; Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khajak, H.G.; Glodek, A.; McKenna, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey, L. Nature 388, 539-547, 1997
 A/Authors: Wallin, B.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Frazer, C. A/Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.
 A/Reference number: A64520; MUID:97394467; PMID:9252185
 A/Accession: C64686
 A/Status: preliminary; nucleic acid sequence not shown; translation not shown
 A/Molecule type: DNA
 A/Residues: 1-228 <TOM>
 A/Cross-references: GB:AE00634; GB:AE00511; NID:G2314489; PIDN:AAD08372.1; PID:G231449
 C/Superfamily: hypothetical protein b2682

Query Match 18.5%; Score 235; DB 1; Length 228;
 Best Local Similarity 33.0%; Pred. No. 9, 2e-13; Indels 14; Gaps 7;
 Matches 70; Conservative 36; Mismatches 92;

Qy 44 LGWYPIGIAFGILVIOGYEWMAAPLPSGLIFAGSTEMVIALVWGAAPLGAIALTLTLLV 103
 Db 20 LGYILMGMTGMLLVQGGYQYKVALFMSLFYVAGAVQFAVITLLSQASLMMVIVSLIV 79
 Qy 104 NFRHVFYAFSPFLHVNKNPIAR-YGVFALIDEAVT--TAARPAQMSAMRLI-SMQIAF 159
 Db 80 NARQTCYALSM-LDRFNKTKMRPLVLAHALTDETFALLNLVAPKGVSSKDFISISLNN 138
 Qy 160 HSFWVFGGLTGVAAIAELIPEIKGLEFALCSLFTVTLTDSCTRKQIPSLLAGSFTIA 219
 Db 139 HSWVIFGSLVSGVSHFSFDTGMEFVMTALFIVLFMEQ--YKRTNHKNAMLGIVIA 195
 Qy 220 LVVIPGQALFPA--LILFGLLTIKYFFLGK 248
 Db 196 VVCL--ALFRTYFLLIALVLMVLAFLMR 224

RESULT 3

F71831
 Hypothetical protein Jhp1251 - Helicobacter pylori (strain J99)
 C/Species: Helicobacter pylori
 A/Variety: strain J99
 C/Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 20-Sep-1999
 C/Accession: F71831
 R/Alt: R.A.; Ling, L.S.L.; Molr, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.R.; Ives, C.; Gibson, R.; Werberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.; Nature 397, 176-180, 1999
 A/Title: Genomic sequence comparison of two unrelated isolates of the human gastric path A/Reference number: A71800; MUID:99120557; PMID:9923682
 A/Accession: F71831
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-228 <ARN>
 A/Cross-references: GB:AE001548; GB:AE00439; NID:G4155845; PIDN:AAD06824.1; PID:G415585
 A/Experimental source: strain J99
 C/Genetics:
 A/Gene: jhp1251
 C/Superfamily: hypothetical protein b2682

Query Match 18.1%; Score 229.5; DB 2; Length 228;
 Best Local Similarity 29.6%; Pred. No. 2, 7e-12; Indels 9; Gaps 5;
 Matches 61; Conservative 41; Mismatches 95;

Qy 44 LGWYPIGIAFGILVIOGYEWMAAPLPSGLIFAGSTEMVIALVWGAAPLGAIALTLTLLV 103
 Db 20 LGYILMGMTGMLLVQGGYQYKVALFMSLFYVAGAVQFAVITLLSQASLMMVIVSLIV 79
 Qy 104 NFRHVFYAFSPFLHVNKNPIAR-YGVFALIDEAVT--TAARPAQMSAMRLISMQIAF 159
 Db 80 NARQTCYALSM-LDRFNKTKMRPLVLAHALTDETFALLNLVAPKGVNETDFMFSISLNN 138
 Qy 160 HSFWVFGGLTGVAAIAELIPEIKGLEFALCSLFTVTLTDSCTRKQIPSLLAGSFTIA 219
 Db 139 HSWVIFGSLVSGVSHFSFDTGMEFVMTALFIVLFMEQYKSNTHKNAML-GIAIAVY 197
 Qy 220 LVVIPGQALF--AALIFGLLTI 242
 Db 198 CLALFRTYFLLIALVLMVLAFLMR 223

RESULT 4

F64041
 Hypothetical protein H11738 - Haemophilus influenzae (strain Rd KM20)
 C/Species: Haemophilus influenzae
 C/Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 28-Jul-2000
 C/Accession: F64041
 R/Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, J.; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.; Glodek, A.; Kelley, J.M.; Weidman, C.; D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhmann, J.L.; Geoghegan, N.S.W. Science 269, 496-512, 1995
 A/Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Frazer, C.M.; Smith, H.O.; Venter, A/Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd. A/Reference number: A64000; MUID:95350630; PMID:7542800
 A/Accession: F64041
 A/Status: nucleic acid sequence not shown; translation not shown
 A/Molecule type: DNA
 A/Residues: 1-244 <TIGR>
 A/Cross-references: GB:U32846; GB:L42023; NID:G3212237; PIDN:AC23382.1; PID:G1574596; 1
 C/Superfamily: hypothetical protein b2682

Query Match 17.9%; Score 227.5; DB 1; Length 244;
 Best Local Similarity 27.5%; Pred. No. 4, 3e-12; Indels 13; Gaps 6;
 Matches 57; Conservative 50; Mismatches 87;

Qy 45 GMYPIGIAFGILVIOGYEWMAAPLPSGLIFAGSTEMVIALVWGAAPLGAIALTLTLLV 103
 Db 30 GFLPLGIVAGVYMAALGPGFLYPTLMLTIVAGSVERIAGALAPSPSI-SVLLITIMI 88
 Qy 104 NFRHVFYAFSPFLHVNKNPIARYSVPAIDEXA--VTAAAPAGMSAMRLISMQ 156
 Db 89 SARQIFYGISMLEKYGHIHCK--RWYLLITLVDEBSFSLNYAKIPHLDRKGVMEFVS 145
 Qy 157 IAFHSYWFVFGGLTGVAAIAELIPEIKGLEFALCSLFTVTLTDS-CTRKQIPSLLAGS 215
 Db 146 LYLHIVYLGAAMKGLPFTVLPFNKGVFEMTMLFVIFPKNMLKQSHSSSLGIGIA 205
 Qy 216 FTIALVIVPGALFPAALIFGLLTI 242
 Db 206 LVFLITIGKEYFLIPTLIGIWMILTM 232

RESULT 5

C85916
 Hypothetical protein 23983 [imported] - Escherichia coli (strain O157:H7, substrain EDL)
 C/Species: Escherichia coli
 C/Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 01-Mar-2002
 C/Accession: C85916
 R/Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew Miller, L.; Grobeck, E.J.; Davis, N.W.; Linn, A.; Dinalanta, E.; Potamousis, K.; Apodaca, Nature 409, 529-533, 2001
 A/Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 16, 2004, 06:51:17 ; Search time 19.5766 Seconds

(without alignments)
1233.313 Million cell updates/sec

Title: US-10-608-504-3

Perfect score: 1269

Sequence: 1 MORTQEIHSLEVSPEKAL.....LIFLGLTRRYFLGKAAR 251

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 78:*

1: PIR1:*

2: PIR2:*

3: PIR3:*

4: PIR4:*

SUMMARIES

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	Score	Query Match Length	ID	Description
1	240	18.9	254 1	branched-chain am
2	235	18.5	228 1	conserved hypothet
3	229.5	18.1	228 1	hypothetical prote
4	227.5	17.9	244 1	hypothetical prote
5	198	15.6	245 2	hypothetical prote
6	198	15.6	245 2	hypothetical prote
7	197	15.5	245 1	hypothetical prote
8	181	14.3	234 2	branched-chain am
9	181	14.3	234 2	branched-chain am
10	181	14.3	239 2	azlC family protei
11	177.5	14.0	218 2	conserved hypothet
12	175	13.8	252 2	hypothetical prote
13	175	13.8	252 2	hypothetical prote
14	175	13.8	252 2	hypothetical prote
15	174.5	13.8	235 2	probable amino act
16	169	13.3	237 2	amino acid permeas
17	166	13.1	235 2	branched-chain am
18	163.5	12.9	219 2	transport protein
19	147	11.6	231 2	probable exported
20	145.5	11.5	240 2	hypothetical prote
21	145.5	11.5	240 2	conserved hypothet
22	145	11.4	219 2	azlC family protei
23	140.5	11.1	219 1	conserved hypothet
24	140	11.0	246 2	branched-chain am
25	137	10.8	235 2	hypothetical prote
26	115	9.1	235 2	azlC family protei
27	113.5	8.9	477 2	azlC-related prote
28	112	8.8	379 2	probable amino ac
29	112	8.8	714545	Na+/H+ antiporter
				probable sugar tra

30	110.5	8.7	372 2	J02135	NADH dehydrogenas
31	110.5	8.7	508 2	AH2688	Na+/H+ antiporter
32	110.5	8.7	508 2	AH2688	Na+/H+ antiporter
33	108.5	8.6	334 2	B96997	probable NADH dehy
34	108	8.5	256 2	T35133	permease (imported
35	107	8.4	399 2	AH2542	hypothetical prote
36	106.5	8.4	487 2	E96782	hypothetical prote
37	103	8.1	408 2	B83333	hypothetical prote
38	103	8.1	408 2	B83333	probable MFS trans
39	102	8.0	401 2	D87889	NADH dehydrogenase
40	101.5	8.0	401 2	C95920	hypothetical membr
41	100	7.9	417 2	T51467	glucose 6 phosphat
42	100	7.9	260 2	E83174	hypothetical prote
43	100	7.9	541 2	AH2909	hypothetical prote
44	99.5	7.8	541 2	F97684	thiamin transport
45	99	7.8	455 2	A75121	hypothetical prote
			233 2	F97382	probable transport

ALIGNMENTS

RESULT 1

branched-chain amino acid transport azlC - Bacillus subtilis

CISpecies: Bacillus subtilis

CIDate: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 21-Jul-2000

CAccession: G69592; T44777

RXnust: F.; Ogawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berter

C.; Bron, S.; Brouillet, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Ch

A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabre, C.; Ferrari, E.

Nature 390, 249-256, 1997

AAuthors: Foulger, D.; Fultz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gall

iesch, U.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.

Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,

Y. M.; Ogawa, K.; Ogiwara, A.; Oudeg, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle

Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadate, Y.; Sato, T.; Scanlon,

AAuthors: Schleich, S.; Schroeter, R.; Scoffone, F.; Seitzinger, J.; Sekowska, A.; Seror

akenuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama

T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasunoto, K.; Yata, K.; Yoshida, Y

AAuthors: Yoshikawa, H.F.; Zundstein, E.; Yoshikawa, H.; Zundstein, E.

A>Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.

A:Reference number: A69580; MID:198044033; PMID:9384377

A:Accession: G69592

A:Statut: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-254 <BFL>

A:Cross-references: GB:299117; GB:AL009126; NID:G2634966; PIDN:CA814612.1; PID:G2635116

A:Experimental source: strain 168

R:Belitsky, B.R.; Gustafson, M.C.U.; Sonenshein, A.L.; von Wachenfeldt, C.

U:Biochem J. 179, 5448-5457, 1997

A>Title: An lrp-like gene of Bacillus subtilis involved in branched-chain amino acid tra

A:Reference number: Z22837; MID:97431495; PMID:9287000

A:Accession: T44777

A:Statut: translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-254 <BFL>

A:Cross-references: EMBL:Y11043; NID:G1926275; PIDN:CA81940.1; PID:G1926281

A:Experimental source: strain 1A1

C:Genetics:

A:Gene: azlC

C:Superfamily: hypothetical protein b2662

Query Match 18.9% Score 240; DB 1; Length 254;

Best Local Similarity 30.5% Pred. No. 3.9e-13;

Matches 62; Conservative 44; Mismatches 89; Indels 8; Gaps 6;

QY 45 GMYPIGIAFLVYQGYEWMAAPLPSGLIPAGSTEMYIALVGA-APLGAIALTTLV 103

DB 42 GFLPLGIAYGIFMHSIGFSAIYPIIMSPMIFAGSMFVAFNPLGAFNMNALFL-TLV 100

QY 104 NFRVVFAPSPFLHVVNPIARFYSFALIDEVAA--TAAPDAG-SAMRISQIAFH 160